THE MECHANISM OF DNA SENSING BY AIM2-LIKE RECEPTORS

By
Seamus R. Morrone

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
July 2016

Abstract

AIM2-Like Receptors (ALRs) are a family of nuclear and cytosolic foreign-DNA sensors consisting of an N-terminal PYD and one or two C-terminal DNA-binding HIN200 domains. AIM2 is a cytosolic sensor which forms a supramolecular structure known as the inflammasome, while IFI16 resides in both the nucleus and cytoplasm, where it can associate with ASC and form its own inflammasome, as well as lead to activation of the interferon pathway by parallel means. Persistent questions in innate immunology remain of how such sensors respond to foreign DNA while remaining silent towards host DNA, as well as their role in activating downstream effectors.

Chapter 2 is devoted to biophysical and biochemical studies of IFI16. Using fluorescence anisotropy, FRET, EMSA, and various mutational studies, it is shown that the PYD of IFI16 plays a positive, cooperative role in DNA binding, allowing it to oligomerize in a length-dependent manner. Disruption of the non-DNA-binding PYD in turn disrupts DNA binding, demonstrating that oligomerization and DNA binding are coupled events.

In Chapter 3, attention is turned to AIM2. An autoinhibition model exists in the literature to explain AIM2 activation, and this model is tested. By fluorescence anisotropy, FRET, EMSA, electron microscopy, and mutagenesis studies, a simpler model which doesn't invoke autoinhibition is put forth to explain the behavior and activation of AIM2. Results also give clues into how AIM2 may then recruit the next member of the downstream pathway, ASC.

Chapter 4 revisits IFI16 to explore an outstanding question regarding the self-vs-nonself problem; namely, do nucleosomes inhibit IFI16 DNA binding and oligomerization? Using time-dependent FRET assays as well as competition anisotropy experiments, it is demonstrated that nucleosomes act as effective barriers to IFI16 DNA binding and oligomerization. The length-dependent rates of FRET signal also support a model in which IFI16 uses one-dimensional diffusion as an efficient means of oligomerization.

Thesis Advisor: Jungsan Sohn
Thesis Committee: L. Mario Amzel, Scott Bailey, Greg Bowman, Cynthia Wolberger

Acknowledgements

The first people I would like to thank are my parents, who have provided me with an enduring desire to understand the world and a work ethic to get the job done. My being here now could not have been fathomable without their help in making me who I am. My brother has played just as large a role as a sibling and best friend. I want to thank my wife, Stefany, who has given the sometimes dark times of graduate school a light at the end of the day. These failures and triumphs have also been shared with my boss and mentor, Jungsan, who has given great guidance over these years. My lab mates have been great company and a good test audience for ideas and experiments. The other members of the department and program have made these years in Baltimore a wonderful experience I'll keep close to my heart. Thank you all for your help and support.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Every living organism encodes some means to protect itself against external threats. From *E. coli* to humans, all are confronted with threats occurring with such regularity that such a mechanism is as essential as the means to propagate the genomic material. Some of these are relatively uncomplicated, such as the restriction/modification enzyme system in various bacteria [1], while some are complex and can even result in changes to the organism's own genome (CRISPR in bacteria, VDJ recombination in white blood cells). These defenses are referred to, collectively, as the immune system. Many vertebrates, and mammals in particular, possess an immune system composed of two branches, the innate and adaptive immune system [2]. The adaptive immune system is optimized for long-term, specific-target defense: by generating antibodies against poliovirus, one acquires immunity to it for many years, but not to other, related, enteroviruses. These antibodies are crafted by a complex process that results in a highly-specific, high affinity interaction between the antibody and one part of the antigen [3]. While this system is fine-tuned to recognize a unique signal in a sea of various other possible signals, a drawback is its speed, or rather, the lack thereof. The time course of an adaptive immune response is on the order of days to weeks [2]. With such a system, an organism would quickly succumb to an invader. Thus, another means must exist to prevent us from disintegrating under the onslaught of enemies.

In contrast to this slow, specialized response to pathogens is the fast, generalized innate immune system. In mammals, this consists of an assortment of tissues like the mucosal epithelia, specialized cell types like natural killer cells, and protein complexes such as the inflammasome [3]. Various types of cells patrol the

body in search of general danger signals, either in the form of pathogen-associated molecular patterns (PAMPs) - those signals arising from pathogens such as structural components of bacteria - or danger-associated molecular patterns (DAMPs), which are signals indicating damage or potential damage to the cell [4] [5]. These guardian cells, as well as other cell types, express various molecular systems to sense these signals. Some of these systems reside on the cell surface to detect external threats, while others, and the focus of this thesis, reside within the cell, either in the nucleus or cytoplasm, to detect internal threats. These internal threats may either be interior damage, such as lysed compartment components, or interior invasion, such as by a virus or bacterium. Upon detection the response can be measured, not in days like the adaptive immune response, but in terms of minutes and hours.

One of the hallmark responses of the innate immune system is inflammation, a term that encompasses a number of actions on the macroscale and microscale. At the visible level, we see swelling and reddening of the infected tissue; were we to turn to the microscope, we would see secretion of signaling molecules, recruitment of guardian cells, silencing of pathogen gene expression, and engulfment and destruction of the pathogens. Inflammation doesn't lead only to the death of the invading cells, but also to our own. Another major outcome is a controlled form of cell death, much like apoptosis, referred to as pyroptosis. This functions to destroy cells that have already become infected, as well as to help spread signals to neighboring cells to assist them in combating the infection. The molecular machine

responsible for turning on this inflammation-associated cell death pathway is called the inflammasome [4] [6].

The canonical pathway for assembly of the inflammasome and its downstream effects is as follows [4]: upon activation by its respective signal (be it DNA, flagellin, urea crystals, or some other PAMP/DAMP), the receptor recruits an adaptor molecule. The adaptor molecule serves as bridge between the sensor and the effector, which in the canonical context is procaspase-1. Procaspase-1 is the zymogen form of caspase-1, a cysteine protease that cleaves after aspartate residues. This binding of procaspase-1 to the adaptor leads to activation and autocleavage into caspase-1, which goes on to activate the interleukins interleukin-1β and interleukin-18 by processing the pro-forms. Interleukins, in turn, are secreted, which then turn on the inflammation pathway by activating NF-κB, ultimately leading to the release of pro-inflammatory cytokines [4].

Thus, the inflammasome consists of, usually, three members: the sensor, the adaptor, and the effector; sometimes the sensor and adaptor may be united in one molecule, or sometimes it may consist of an extra enhancer member; sometimes the effector may be a different caspase. Usually, however, the adaptor and effector remain the same, while the sensor may be switched out, allowing for the detection of a different signal. So one may consider "the" inflammasome as, really, a family of inflammasomes, which may be distinguished by referring to a specific one as "the AIM2 inflammasome" or "the NLRP3 inflammasome," the qualifier coming from the name of the sensor.

The nature of the signal is of interest to us for many reasons. At the basic level, we may ask: how do these sensors, which are germ-line encoded and thus, in contrast to antibodies, are rather inflexible in terms of signal recognition, recognize such a wide variety of different pathogens? Obviously, differences between the pathogen and host should be exploited. Luckily, there are many characteristics that differ between, for instance, a human cell and a bacterial cell; and which, therefore, are possible signals to be used as a warning. The flagellin of a mobile bacterium is a prime example; and so it is found to be a signal for the Nod-Like Receptor (NLR) family [4]. The presence of a cell wall in Gram-negative bacteria may serve as another signal – and indeed it is [4]. However, there are many organisms which don't possess these structures and a different, more general signal must be used, and this differentiation begins to break down. As we draw close to signals that are universally shared amongst all organisms – DNA, for instance – we begin to see the thin ice our defense system treads in this battle. Yet even here we are not totally lost. Signature motifs may occur in some organisms which do not occur in ours: unmethylated CG dinucleotide (CpG) motifs are one example present in bacteria and DNA viruses but rare in mammals, and the Toll-Like Receptor family member TLR9 is one that recognizes these [7] [8]. AT-rich sequences (a characteristic of some pathogens) are recognized by Pol III. Yet what protection does our defense system provide against an enemy which possesses none of these?

Many times, compartmentalization provides the ability to distinguish self from nonself [3]. DNA ought, under normal circumstances in our cells, to be found in the nucleus. A sensor patrolling the cytoplasm would then detect any DNA, which

presumably would be there only if foreign DNA had been inserted into the cytoplasm or the integrity of the nucleus had been challenged. If these sensors resided solely in the cytoplasm, then perhaps this question of self-versus-nonself might not be so interesting – but, in fact, many of them have been reported to localize in the nucleus. What is the mechanism of such a defense? How does it work in one instance (against the invader) and not in another (against the self)? Understanding how they work would provide unique insights into this central question of innate immunity, but might also provide fodder for how these sensors may misbehave, giving rise to such diseases as Sjögren's syndrome, systemic lupus erythematosus, and psoriasis [9] [4]. The focus of the thesis work presented here has been on a specific group of these foreign DNA sensors.

These sensors belong to a family of proteins that has been known by various names throughout its history. It has been called the IFI protein family [10]; p200 proteins [11]; PYHIN proteins [12], and, most recently, AIM2-Like Receptors [13]. What this plethora of names reveals is our attempt to encapsulate the essence of the family: shall we group them by their homology to a given member (AIM2-Like Receptors), or by their shared inducibility (IFI protein family), or perhaps by one of their characteristic domains (HIN-200 proteins) – perhaps by both of their characteristic domains (PYHIN proteins)? Each of these appellations provides insights into this family of sensors. In terms of expression, it has been found they are induced by type I and type II interferons via consensus IFN-responsive elements (ISRE) [14] [15]. The recent naming suggests their similarity to what might be considered the quintessential member, but probably is more a reflection of the

community's eagerness to have a unifying naming convention (witness the Toll-Like Receptors [TLR], the C-type lectin receptors [CLR], the RIG-I-like receptors [RLR], and the NOD-like receptors [NLR]) [16]. Their similarity to each other, however, is not forcing a square protein in a circular family, but is very real and probably arose from gene duplication events. Thus, these family members consist of the same general domain organization.

The C-terminal domain consists of the portion that recognizes DNA - the HIN200 domain [17] [18]. This acronym is short for hematopoietic, interferon-inducible nuclear domain of 200 amino acids, of which two of these modifiers are not rigidly followed. While they are expressed in hematopoietic cells, they are not exclusively expressed there (AIM2 is found in keratinocytes [9]); and while they are found in many proteins localized to the nucleus, they are not exclusively found in those (once again, AIM2 being a prime example [9]). It consists of two subdomains that are insoluble when individually expressed [19], the OB-fold (oligosaccharide/oligonucleotide binding fold), which is the core of what recognizes DNA. This recognition is mediated almost exclusively by electrostatics: the negatively-charged phosphate backbone of DNA interactions with a positively-charged, basic patch of lysines and arginines in the HIN200 domain [20]. This electrostatics-based recognition ensures a nonspecific binding; its ability to distinguish DNA from RNA, or single-stranded from double-stranded, is less well-understood. It may be that residues inserted into the major groove of DNA grant an ability to distinguish DNA from RNA [9], or the width of the binding pocket of the protein restricts the ligand to double-stranded DNA. Reports of isolated HIN200

domains binding RNA or single-stranded DNA are in the literature [21], but have not been found to hold in the context of the wild-type protein [18]. Little conformational change occurs upon binding to DNA, which hints that this domain must not function as the sensor in itself; there must be something else to propagate the signal. This leads us to the next domain.

In keeping with this proliferation of names, the N-terminal domain has gone by various names in the course of its career: called the pyrin domain in this work, it has also been called the DAPIN domain or the PAAD domain. The pyrin domain (PYD) belongs to a much larger superclass called the death domain (DD) superfamily. This chilling moniker arises from the frequency of these domains in proteins involved in cell death pathways. It's a deceptively simple piece of protein, consisting of six alpha helices bundled together into a little globule, a form characteristic of most DDs. This domain was first characterized as belonging to the gene product, called pyrin, responsible for Familial Mediterranean Fever, which is characterized by bouts of inflammation as well as amyloidosis. They are also only found at the N-terminus of proteins [22], indicating this positioning may be important to its function. Inevitably in perusing the literature on this domain, one encounters the phrase "homotypic interaction" in some form or another [8] [23] [11] [9] [24]. This cryptic modifier can be expanded thusly: a protein containing a PYD generally can interact with another protein containing a PYD. There are, of course, exceptions and footnotes to this rule, but by and large as a rule of thumb it suffices. This is the basis for its signaling role, and coupled with the HIN200 domain it appears we have constructed a minimal sensor.

What has emerged in the last few years is an appreciation of the propensity of the PYD (and, incidentally, many other DDs) to form long, thin filaments at high concentrations, which appear to be essential for their ability to activate the next components in the pathway [25]. Many of these filaments shared conserved interfaces that allow for templating of the next component of the pathway. The helical filament is believed to provide an advantage of recruiting components of the signaling pathway by providing a larger recruitment surface, as well as providing a means to generate a distinct on/off mechanism.

The ALRs work by activating two seemingly mutually exclusive pathways in the immune system: either maturation of the interleukins via ASC to caspase, or the STING-TBK1-IRF3 axis [26] leading to IFN-β production. IFN-activating pathways are not affected by knocking out ASC [18], demonstrating its independence. Yet the role of a given ALR is not limited to one pathway, as some members have been reported to activate different pathways depending on the context [27].

The focus of this thesis work has rested on two members of the ALRs, AIM2 and IFI16, as well as their shared adapter protein ASC. These will now be introduced in a more detailed light for elucidation of the coming pages of experiments and findings.

Interferon-inducible protein 16 (IFI16) [28] has been implicated in diverse pathways of the organism, ranging from tumor suppression [19] to inflammation. IFI16 is expressed in myeloid cells and fibroblasts as three different isoforms (A, B, and C), of which the B form appears to be the most relevant in terms of expression level and distribution [29] [11]. These isoforms differ only in the length of linkers

between the three domains of the protein. At the N-terminus, of course, is the PYD. This is followed by a long, unstructured linker region containing nuclear-localization signals, as well as some bona fide and putative post-translational modification sites [30]. While IFI16 is found predominantly in the nucleus, it has also been isolated in the cytoplasm. Unusual in the human repertoire, IFI16 possesses two HIN200 domains, both of which have been shown to bind DNA.

It wasn't until 2010 that the important role IFI16 plays in recognizing and responding to foreign DNA was truly appreciated [18] . Until that time, it had been characterized offhandedly in conjunction with DNA, such as in the DNA-repair complex BASC, as a transcriptional repressor, or as a partner that enhances the binding of p53. This binding of foreign DNA induces the recruitment of STING, which then activates IRF3 and NF-κB, leading to the expression of IFN-β. There appears to be a strong length-dependency in vivo, as DNA of less than 70bp do not generate a strong production of IFN-β [18]. This binding appears to be, as in the case of the isolated HIN200 domains, nonspecific, but cooperative [18].

IFI16 appears to be special not just due to its extra HIN200 domain, but also by the fact that it acts as a foreign DNA sensor in both the cytoplasm [18], where it activates the IFN-β pathway, and the nucleus, where it activates the inflammasome pathway [27]. This shuttling between the two subcellular compartments is due to the NLS located in the linker region between the PYD and the first HIN200 domain, which upon acetylation prevents IFI16 from being transported from within the cytoplasm to the nucleus [30]. While its cytoplasmic sensing of foreign DNA presents no conceptual difficulties, this ability to act as a DNA sensor within the

nucleus, where presumably the concentration of host DNA far exceeds that of invader DNA, is not well understood. Evidence of IFI16 as a viral gene silencer and its promotion of heterochromatin formation suggesting that chromatinazation plays an inhibitory role in this sensing have been put forth [31]. This is an attractive hypothesis, as histone-associated DNA may be effectively compartmentalized away from IFI16, and linking DNA between histones is between 20-30bp, below the length seen to induce robust IFN-β production. However, no in vitro studies have yet to be performed to directly test this hypothesis, as well as to confirm the length-dependency observed in vivo.

Absent in Melanoma 2 (AIM2) was first isolated and characterized by Trent and colleagues [32] (Ray et al, 1996). By Northern blot analysis, they found it significantly expressed in spleen, small intestine, and peripheral blood leukocytes [32]. As with other ALR family members, AIM2's expression following IFN-β induction, as measured via RNA transcripts, was significantly increased in the first hour [32]. Of note, AIM2 appears to be a cytosolic DNA sensor, while the other members reside in the nucleus [23]. There it interacts with the PYD of ASC to form the AIM2 inflammasome [23] [33]. There are two isoforms of AIM2 [9]. Like IFI16, AIM2 shows a strong length-dependency, requiring around 80bp to generate a robust production of IL-1β [20], but such a response can't be divined from the flavorless, phosphate-backbone-binding cartoon. So whence this dependency?

A model of AIM2 activation was put forth by Xiao and colleagues after the release of a paper of the crystal structures of the HIN domain bound to DNA [20], as well as a following article of the crystal structure of MBP-tagged AIM2 PYD [17].

Through molecular docking they were able to show the PYD fits snugly into the DNA-binding pocket of the HIN200 domain, the basic patch of the HIN200 interacting with an acidic patch on PYD. Through the use of fluorescence anisotropy assays of MBP-tagged versions of the full-length protein and the HIN200 domain, they showed that the HIN200 domain appeared to bind with better affinity than the full length. These lines of evidence suggested to them the following model. In the resting state, the PYD of AIM2 is folded back and interacting with the DNA-binding pocket of the HIN200 domain. This prevents the PYD from spuriously associating with its binding partners. DNA acts as a competitor for this state, displacing the PYD as it binds to the HIN200 domain and activating AIM2 so that it is free to bind to ASC.

Questions with regard to this model arise, however. As the DNA-binding footprint of the HIN200 domain is only 8-bp as shown in the crystal structure, from whence does the length-dependent IL-1b secretion response arise? Would not all DNA of length exceeding the footprint allow for activation of AIM2? While an autoinhibition mechanism for ALRs is attractive, as it provides unification with other innate immune receptors that operate in such a fashion, such as NLRs and RLRs, it may not generalize to other members. In particular, IFI16 has been shown to behave in solution as an extended rod, inconsistent with an autoinhibition model [19]. Recently, it has come to light that the AIM2 PYD behaves very similarly to other DDs. At high concentrations, it appears to form filaments. A paper reporting the crystal structure of a PYD variant in which a critical residue for oligomerization had been mutated revealed many interesting discrepencies between it and the MBP-

tagged crystal structure. Importantly, it showed that an important interface for oligomerization was masked in the MBP-tagged version of the PYD, suggesting that the tag interferes with the native behavior. This was borne out by the fact that the MBP-tagged version remains soluble and monomeric, while the untagged AIM2-PYD readily crashes out of solution and forms large filaments. The paper also demonstrated a number of mutations in the oligomerization interface which could disrupt polymerization in vivo. The role of this polymerization in the autoinhibition model has yet to be addressed.

The last protein that makes up the list explored by this thesis is ASC, referred to above as the adaptor protein. ASC is an acronym for "apoptosis-associated speck-like protein containing a caspase activating and recruitment domain," of which this meaning will be dissected. Caspase activating and recruitment domain (CARD) is another DD that, in the context of ASC, assists in the conversion of procaspase-1 to caspase-1 by binding to the CARD of procaspase-1. ASC is a small, two-domain cytosolic protein, consisting of an N-terminal PYD and the C-terminal CARD. As ASC is an essential component of the inflammasomes, its role in apoptosis, and thus the first part of its name, should be apparent. As the adaptor molecule of the inflammasome, ASC is able to interaction with many of the sensors via its PYD. This includes the ALRs associated with the inflammasome, AIM2 and IFI16, as well as the other major inflammasome sensors, the NLRs, such as NLRP3. Finally, upon activation, in vivo studies have found ASC to form high molecular weight, speck-like punctates in the cell [34]. In vitro, the isolated PYD and CARD of ASC have been shown to undergo polymerization, which can be disrupted by select mutations.

Interestingly, ASC has been shown to behave in a manner similar to prions, in which ASC forms high-molecular weight, detergent-resistant filaments, which in the extracellular space can become internalized by nearby cells and convert their endogenous, monomeric ASC into filaments [35]. This study provides insight into how the inflammation process may be propagated across cells by using components of the detection pathway rather than signaling molecules such as interleukins. Importantly, it suggests a switch-like mechanism for the pathway, in which a small pool of activated molecules may then act as the nucleating point for inactivated members.

Wu and colleagues [33] showed that the presence of AIM2, either as the PYD alone or as the full-length in complex with DNA, leads to an enhancement of the polymerization of ASC. This is consistent with another report demonstrating that AIM2 enhances the prion-like activities of ASC in vivo [35]. By using gold-labeled AIM2-PYD, they were able to show AIM2 localized to one end of the ASC filament, indicating that indeed a sub-stoichiometric amount of AIM2 allows for nucleation of ASC. They also provided a high-resolution cryo-electron microscopy map of the filament, showing it as a helical assembly with specific symmetry parameters, and verified their model by select mutagenesis. It has been proposed that the AIM2-PYD filament will match the symmetry of the ASC filament, thus acting as the template for ASC polymerization [36], but the structure and symmetry elements of AIM2 have not yet been determined to allow for stronger support of this idea. If this were to be correct, then activated AIM2 may in fact be the nucleator of the prion-like activity observed for ASC.

Combining the above outline of the players of inflammation, we can revise our simple model in a few places of inflammasome activation. When AIM2 in the cytosol or IFI16 in the nucleus detects foreign DNA (how it can distinguish foreign from host still remains to be seen) it becomes activated, which in turn leads to recruitment of ASC via PYD-PYD interactions. ASC binding to AIM2 leads to polymerization. This filament then leads to binding of procaspase-1 via CARD-CARD interactions. Due to clustering on the filament of ASC, a local high concentration allows self-activation of procaspase-1 by autocleavage, generating mature caspase-1.

While the broad details of the pathway appear to be articulated, gaps exist which prevent a smooth link between steps. Two major outstanding questions of this pathway, and the focus of this thesis work, are 1) how do ALRs detect foreign DNA while at the same time remaining silent to host DNA, and 2) how is this detection of signal propagated to the next component of the pathway, namely ASC? In regards to the first question, the current model suggests an autoinhibited state that should equally respond to both foreign and host DNA given the correct conditions [17]. This question is particularly germane to IFI16, which resides in the nucleus and for which an autoinhibited state does not appear to exist [19]. With regard to the second question, as current research and opinion in the field leans towards a two-state, digital model, in which the cell is distinctly either not in the inflammation-signaling state, or is distinctly in such a state, the mechanism of ALR activation ought to have such an outcome built into it. The following chapters will explore these two questions.

References

[1] ER Westra et al., "The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity," *Annu Rev Genet*, vol. 46, pp. 311-39, 2012.

[2] AK Abbas, AH Lichtman, and JS Pober, "Cellular and Molecular Immunology," 1991.

[3] R Medzhitov, "Recognition of microorganisms and activation of the immune response," *Nature*, vol. 449, pp. 819-826, 2007.

[4] F Martinon, A Mayor, and J Tschopp, "The inflammasomes: guardians of the body," *Annu Rev Immunol*, vol. 27, pp. 229-65, 2009.

[5] JC Kagan, VG Magupalli, and H Wu, "SMOCs: supramolecular organizing centres that control innate immunity," *Nat Rev Immunol*, vol. 14, pp. 821-6, 2014.

[6] Daniel A. Muruve et al., "The inflammasome recognizes cytosolic microbial and host DNA and triggers an innate immune response ," *Nature*, pp. 103-107, 2008.

[7] G Barber, "Cytoplasmic DNA innate immune pathways," *Immunol Rev*, vol. 243, pp. 99-108, 2011.

[8] Taner Cavlar, Andrea Ablasser, and Veit Hornung, "Induction of type I IFNs by intracellular DNA-sensing pathways," *Immunology & Cell Biology*, pp. 474-482, 2012.

[9] Neil Shaw and Zhi-Jie Liu, "Role of the HIN Domain in Regulation of Innate Immune Responses," *Molecular and Cellular Biology*, pp. 2-15, 2014.

[10] RW Johnstone, JA Kerry, and JA Trapani, "The human interferon-inducible protein, IFI 16, is a repressor of transcription," *J Biol Chem*, vol. 273, pp. 17171-7, 1998.

[11] Sudhakar Veeranki and Divaker Choubey, "Interferon-inducible p200-family protein IFI16, an innate immune sensor for cytosolic and nuclear double-stranded DNA: Regulation of subcellular localization ," *Molecular Immunology*, pp. 1-5, 2011.

[12] SA Schattgen and KA Fitzgerald, "The PYHIN protein family as mediators of host defenses," *Immunol Rev*, vol. 243, pp. 109-18, 2011.

[13] A Lu et al., "Plasticity in PYD assembly revealed by cryo-EM structure of the PYD filament of AIM2," *Cell Discov*, vol. 1, p. 15013, 2015.

[14] S Kumar, H Ingle, D Vijaya, R Prasad, and H Kumar, "Recognition of bacterial infection by innate immune sensors," *Critical Reviews in Microbiology*, vol. 39, pp. 229-46, 2013.

[15] D Choubey et al., "Interferons as gene activators: indications for repeated gene duplication during the evolution of a cluster of interferon-activatable genes on murine chromosome 1," *J Biol Chem*, vol. 264, pp. 17182-9, 1989.

[16] Ji-Hyun Jang et al., "An Overview of Pathogen Recognition Receptors for Innate Immunity in Dental Pulp," *Mediators of Inflammation*, pp. 1-12, 2015.

[17] Tengchuan Jin, Andrew Perry, Patrick Smith, Jiansheng Jiang, and T. Sam Xiao, "Structure of the AIM2 pyrin domain provides insights into the mechanisms of AIM2 autoinhibition and inflammasome assembly ," *Journal of Biological*

Chemistry, pp. 1-22, 2013.

[18] Leonie Unterholzner et al., "IFI16 is an innate immune sensor for intracellular DNA ," *Nature Immunology*, pp. 997-1005, 2010.

[19] Jack C.C. Liao et al., "Interferon-Inducible Protein 16: Insight into the Interaction with Tumor Suppressor p53 ," *Structure*, pp. 418-429, 2011.

[20] Tengchuan Jin et al., "Structures of the HIN Domain:DNA Complexes Reveal Ligand Binding and Activation Mechanisms of the AIM2 Inflammasome and IFI16 Receptor ," *Immunity*, pp. 1-11, 2012.

[21] Yan Hongyue et al., "RPA nucleic acid-binding properties of IFI16-HIN200," *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, pp. 1087-1097, 2008.

[22] Lan Hoang Chu, Anu Gangopadhyay, Andrea Dorfleutner, and Christian Stehlik, "An updated view on the structure and function of PYRIN domains," *Apoptosis*, pp. 157-173, 2014.

[23] Veit Hornung et al., "AIM2 recognizes cytosolic dsDNA and forms a caspase-1 activating inflammasome with ASC ," *Nature*, pp. 514-518, 2009.

[24] Qian Yin et al., "Molecular Mechanism for p202-Mediated Specific Inhibition of AIM2 Inflammasome Activation ," *Cell Reports*, pp. 327-339, 2013.

[25] Ryan Ferrao and Hao Wu, "Helical assembly in the death domain (DD) superfamily ," *Current Opinion in Structural Biology*, pp. 241-247, 2012.

[26] RL Brunette et al., "Extensive evolutionary and functional diversity among

mammalian AIM2-like receptors," *J Exp Med*, vol. 209, no. 11, pp. 1969-83, 2012.

[27] Nagaraj Kerur et al., "IFI16 Acts as a Nuclear Pathogen Sensor to Induce the Inflammasome in Response to Kaposi Sarcoma-Associated Herpesvirus Infection ," *Cell Host & Microbe*, pp. 363-375, 2011.

[28] Joseph A. Trapani et al., "A novel gene constitutively expressed in human lymphoid cells is inducible with interferon-3, in myeloid cells ," *Immunogenetics*, pp. 369-376, 1992.

[29] Ricky W. Johnstone, Michael H. Kershaw, and Joseph A. Trapani, "Isotypic Variants of the Interferon-Inducible Transcriptional Repressor IFI 16 Arise through Differential mRNA Splicing ," *Biochemistry*, pp. 11924-11931, 1998.

[30] Tuo Li, Benjamin A. Diner, Jin Chen, and Ileana M. Cristea, "Acetylation modulates cellular distribution and DNA sensing ability of interferon-inducible protein IFI16 ," *Proceedings of the National Academy of Sciences, USA*, pp. 1-6, 2012.

[31] Megan H. Orzalli, Sara E. Conwell, Christian Berrios, James A. DeCaprio, and David M. Knipe, "Nuclear interferon-inducible protein 16 promotes silencing of herpesviral and transfected DNA ," *Proceedings of the National Academy of Sciences USA*, pp. 1-10, 2013.

[32] Katherine L. DeYoung et al., "Cloning a novel member of the human interferon-inducible gene family associated with control of tumorigenicity in a model of human melanoma ," *Oncogene*, pp. 453-457, 1997.

[33] Alvin Lu et al., "Unified Polymerization Mechanism for the Assembly of ASC-Dependent Inflammasomes ," *Cell*, pp. 1193-1206, 2014.

[34] Junya Masumoto et al., "ASC, a Novel 22-kDa Protein, Aggregates during Apoptosis of Human Promyelocytic Leukemia HL-60 Cells," *The Journal of Biological Chemistry*, pp. 33835-33838, 1999.

[35] Xin Cai et al., "Prion-like Polymerization Underlies Signal Transduction in Antiviral Immune Defense and Inflammasome Activation ," *Cell*, pp. 1207-1222, 2014.

[36] Arthur V. Hauenstein, Liman Zhang, and Hao Wu, "The hierarchical structural architecture of inflammasomes, supramolecular inflammatory machines ," *Current Opinion in Structural Biology*, pp. 75-83, 2014.

Chapter 2

IFI16 Cooperatively Assembles Into Filaments on dsDNA: Insights Into a Host Defense Strategy

Reproduced in part from:

Morrone SR, Wang T, Constantoulakis LM, Hooy RM, Delannoy MJ, Sohn J. (2014) Cooperative assembly of IFI16 filaments on dsDNA provides insights into host defense strategy. *Proc Natl Sci U S A*. 111(1):E62-71.

2.1 Introduction

The mammalian innate immune system's ability to recognize intracellular foreign DNA is a widely conserved defense mechanism to detect and respond to invading pathogens [1] [2]. Such a universal molecule as a major danger signal for detecting pathogens presents dilemmas that must be stringently and efficiently overcome, yet only a few factors are known which allows the host innate immune system to selectively engage foreign DNA (nonself-DNA) while minimizing interactions with the host DNA (self-DNA); such factors include compartmentalization of the cell and the size of foreign DNA. These features, however, only raise more questions than provide answers. For example, the footprints of intracellular DNA receptors usually fall below 20 bp, and yet a long foreign DNA fragment [e.g., poly(dA:dT); ≥1,000 bp] is required to induce a robust innate immune response even in a normally DNA-free environment like the cytoplasm [1] [2]. On the other hand, foreign DNA-sensing pathways also exist in the host nucleus in which DNA receptors must not respond to abundant self-DNA to prevent spurious activities [1] [2]. Indeed, one of the major unresolved questions in understanding the DNA-sensing pathways of mammals is whether the host intracellular DNA receptors have any capacity to distinguish self- from nonself-DNA at the molecular level [2].

Human IFN inducible protein-16 (IFI16) is an intracellular DNA receptor of innate immunity that belongs to the family of absent-in-melanoma-2 (AIM2)-like receptors (ALRs) [1] [2] [3] [4]. IFI16 senses DNA from invading pathogens in both the nucleus and cytoplasm [e.g., vaccinia virus (VACV) and herpes simplex virus-1

(HSV-1)] and plays critical roles in production and maturation of major proinflammatory cytokines such as IFN-β and interleukin-1β. Notably, IFI16 is the only known host DNA sensor that operates in the nucleus, which is conventionally thought to be off limits to DNA receptors due to the abundant self-DNA. In addition to its role in defense against foreign DNA, the aberrant activity of IFI16 is also associated with several autoimmune disorders, such as systemic lupus erythematosus and Sjögren syndrome. IFI16 is involved in the onset and progression of these diseases not only by generating abnormally high levels of the proinflammatory cytokines but also by being recognized as an autoantigen by the host adaptive immune system [4] [5] [6] [7] [8].

Despite its implication to host defense and autoimmunity, little is known about molecular mechanisms that underlie IFI16. Indeed, there are three major unresolved questions regarding its behavior *in vivo*. First, as observed from several other DNA-sensing pathways in mammals, it is not known why long dsDNA (>60 bp) is required to induce IFI16 activity [3] [9] [10]. Second, why IFI16 does not bind self-DNA in the nucleus is not known [11] [12] [13] [14]. Third, it is not known how IFI16 can selectively assemble into large signaling foci (e.g., inflammasome) with nonself-DNA [11] [12] [13] [14].

Although mechanisms of many proteins can be inferred from their structures, the structure of IFI16 does not provide clear answers to the above questions. IFI16 is composed of two signature ALR domains, namely one N-terminal pyrin domain (PYD) ($IFI16^{PYD}$) followed by two HIN200 domains ($IFI16^{HinA}$ and $IFI16^{HinB}$). The HIN200 domains nonspecifically bind various ss- and dsDNA

fragments on the phosphate backbone via electrostatic interactions with a footprint of eight to nine bases, thus confirming that DNA sequence (e.g., the CpG island) is not a recognition element [9] [15] [16] [17]. On the other hand, the PYD is a homotypic protein-protein interaction domain whose function is thought to be limited to recruiting downstream effectors [3] [9] [10].

Based on crystallographic and binding studies of related AIM2 and the isolated HIN200 domains of IFI16, a general mechanistic model for ALR activation was recently proposed [9] [18]. In this model, ALRs assume an autoinhibited conformation in which the PYD blocks the DNA-binding surface of the HIN200 domain. Upon encountering foreign dsDNA, the PYD is displaced and interacts with its downstream partners. A long dsDNA fragment is used as a pseudooligomerization platform, because the HIN200 domain independently binds to either strand of dsDNA with a footprint of about eight bases. The resulting ALR·DNA complex thus assumes a configuration similar to noninteracting beads on a string. Longer DNA fragments simply accommodate more ALRs, allowing the ALR pseudooligomers to recruit a greater number of downstream effectors.

However, these otherwise seminal studies [9] [18] do not explain any of the three questions regarding IFI16 behavior. First, it provides no rationale for the DNA length-dependent responses observed *in vivo*. By this model, IFI16 would bind equally well to either strand of any dsDNA exceeding its footprint. Thus, any dsDNA long enough to promote minimal IFI16 oligomer for downstream effector interaction (e.g., dimer for procaspase-1 activation; 10 bp) would produce similar responses at the same mass concentration. Second, this model provides no insight

into how IFI16 suppresses its interaction with self-DNA in the nucleus. For instance, AIM2 has only one HIN200 domain and is exclusively localized in the cytoplasm [19]. By contrast, IFI16 has two HIN200 domains that can bind either ss- or dsDNA and assumes an uninhibited open conformation, and its footprint falls within the range of the exposed dsDNA linker between host nucleosomes (10–20 bp) or transcription bubbles (~17 bases) [12] [20] [21] [22]. Indeed, the proposed autoinhibitory mechanism would be effective only in an intrinsically DNA-free environment like the cytoplasm but not in the nucleus, where abundant self-DNA would easily displace IFI16^{PYD} from either HIN200 domain. Finally, it does not explain how IFI16 can selectively colocalize with foreign DNA and assemble into large signaling foci. The total amount of DNA under normal or infectious conditions easily exceeds that of IFI16. Thus, by this model, IFI16 would be scattered on foreign and self-DNA instead of selectively forming signaling foci on nonself-DNA (e.g., in principle, one genome of HSV-1 contains more than 30,000 binding sites).

Here, by using quantitative assays and electron microscopy, I found that cooperatively assembling into filamentous oligomers on dsDNA may underlie the observed behaviors of IFI16 *in vivo*. IFI16 binds dsDNA in a nonlinear length-dependent manner and forms oligomers that are clearly different from entities resembling beads on a string. The isolated HIN200 domains of IFI16 do not oligomerize and thus engage dsDNA with weak affinity. However, in contrast to the PYD of AIM2, IFI16^{PYD} plays an unexpected positive role in DNA binding as it drives the filament assembly. The surface residues that are important for dsDNA binding in IFI16^{PYD} are highly conserved, indicating that other ALRs use the same strategy.

These results reveal that IFI16 is poised to rapidly engage even a single piece of long nonself-DNA with a switch-like mechanism (e.g., HSV-1 genome), while not responding to much higher concentrations of short self-DNA exposed under normal conditions (e.g., nucleosomal linker dsDNA).

2.2 Results

Assembling signal·receptor complexes is a critical regulatory step in virtually all known signaling pathways, because this first step commits the rest of the cascade in a switch-like mechanism [23] [24]. I reasoned that it is unlikely that regulation of IFI16 pathways is only accomplished by yet unknown downstream events. The currently prevailing mechanistic model for IFI16 was established by studying its isolated HIN200 domains, which nonspecifically bind any DNA exceeding their footprints equally well [3] [9] [15]. In contrast, studying the full-length protein might better address the DNA-sensing strategy of IFI16. Thus, for this work, I first successfully generated full-length and the functional domains of IFI16 without solubility tags in high purity (**Figure 2.1A–C**) and also developed various *in vitro* assays to investigate how IFI16 engages DNA.

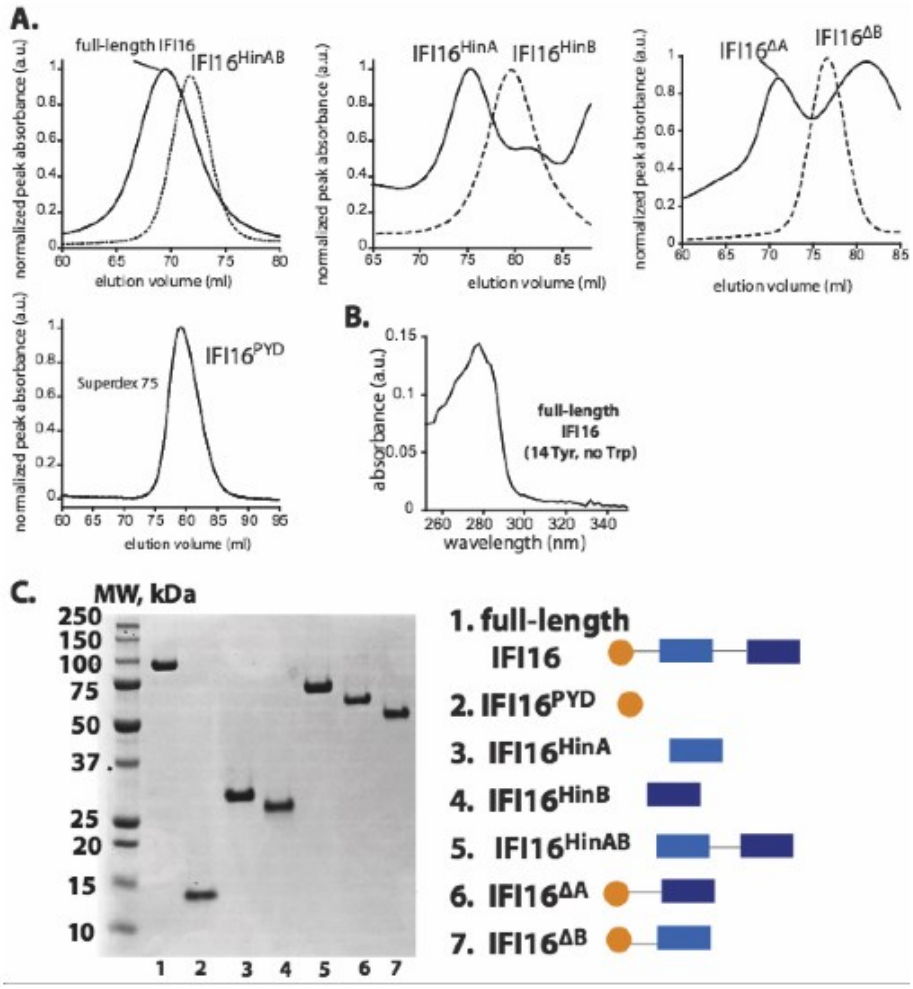

Figure 2.1: Purification of IFI16 constructs used in the study. (A) Sample images from size-exclusion chromatography of IFI16 constructs used. IFI16^{PYD} was purified using Superdex-75 16/600, while all the other constructs were purified using Superdex-200 16/600. All elution volumes correspond to monomers according to the manufaturer's guidelines (GE Healthcare). (B) A Sample UV-visible absorbance profile of full-length IFI16. (C) An SDS-PAGE image of IFI16 constructs used in this study. The gel was stained using Coomassie blue.

2.2.1 IFI16 Binds dsDNA in a Length-Dependent, Cooperative Manner

The DNA length-dependent responses of IFI16 observed *in vivo* may arise directly from differences in binding affinity for different-length DNA ligands but not from an as yet unidentified regulatory mechanism mediated by downstream effectors. To test this, I developed fluorescence-anisotropy (FA) binding assays using various fluorescein-amidite (FAM)-labeled DNA fragments. Monitoring FA of FAM-dsVACV72 (a 72-bp fragment derived from VACV [3]; **Table 2.1**) with increasing concentrations of IFI16 resulted in a sigmoidal binding isotherm which a Hill equation could best fit [an apparent binding constant (K_{Dapp}) of 65 ± 19 nM and a Hill coefficient of 1.9 ± 0.2; **Figure 2.2A**]. This result indicated that full-length IFI16, unlike its isolated HIN200 domains [3] [9] [15], binds dsDNA in a cooperative manner. I then tested the binding of shorter dsDNA variants of VACV and HSV fragments that were used in the previous *in vivo* study [3]. In contrast to predictions from the previously proposed model [9], I found that the binding affinity decreased with decreasing length of the dsDNA fragments independent of sequence (**Figure 2.2B**, **Table 2.1**, and **Table 2.2**). Importantly, the calculated Hill coefficients also decreased with shorter dsDNA (e.g., **Figure 2.2B,C**), suggesting that an absence of cooperativity results in weaker affinity (see also **Table 2.2**).

A possible underlying mechanism for the cooperative length-dependent binding is that IFI16 assembles into a distinct oligomer on dsDNA, because such a protein cluster would result in a more stable complex than the previously proposed noninteracting beads on a string [9]. To test this idea, I first determined the stoichiometry between IFI16 and dsDNA ligands by performing FA experiments

using FAM-dsDNA concentrations at least six times higher than their respective K_{Dapp} values (e.g., **Figure 2.2D**). The plot of dsDNA length vs. the number of bound molecules revealed that about 15 bp are required to accommodate one full-length IFI16 (**Figure 2.2E**), which agrees with a previous report suggesting that one HIN200 domain takes up about eight to nine bases. Notably, because no significant binding was observed from FAM-dsHSV15, these results also suggest that oligomeric IFI16 is required for tight binding (see also below). To confirm the length-dependent binding, I normalized the K_{Dapp} of each fragment to the number of available binding sites on each dsDNA fragment (e.g., four for dsHSV60). Plotting normalized K_{Dapp} vs. dsDNA-length revealed that the binding affinity of IFI16 still increases with increasing DNA length (**Figure 2.2F**), supporting the idea that IFI16 cooperatively oligomerizes on dsDNA in a length-dependent manner. Importantly, these results correlate with the previously reported dsDNA length-dependent responses *in vivo* [3] and thus suggest that cooperatively assembling stable IFI16·dsDNA complexes via oligomerization is critical for regulating its cellular activity.

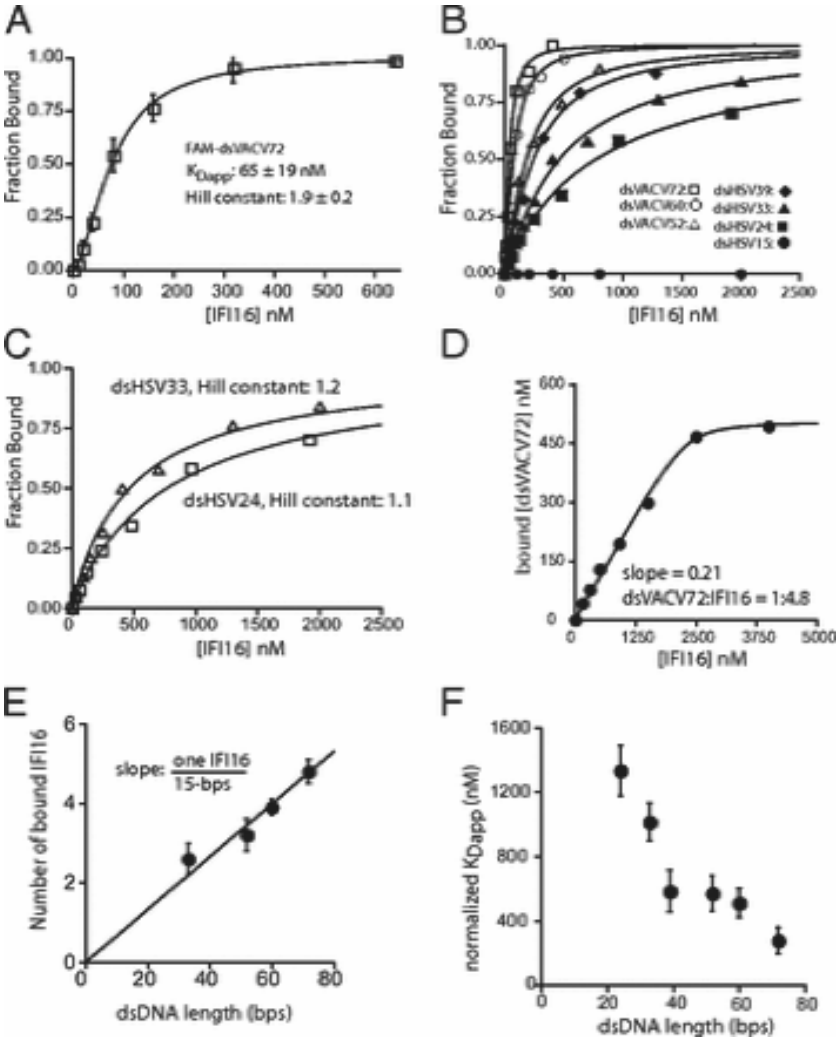

Figure 2.2: IFI16 cooperatively binds dsDNA in a length dependent manner. (A) Binding of full-length IFI16 to FAM-dsVACV72 (2.5 nM) was monitored by changes in FA. The fraction bound was calculated as $(A - A_0)/(A_{max} - A_0)$, where A_0 is the anisotropy of the free DNA and Amax is the anisotropy of the saturating protein. The lines are fits to a Hill form of binding equation (number of independent experiments n = 3). (B) Binding of IFI16 to each FAM-dsDNA was determined by FA. The lines are fits to a Hill equation: fraction bound = $1/(1 + K_{Dapp}/[IFI16])^{Hill\ Coefficient}$). The determined values are listed in Table 2.2. (C) The data for dsHSV33 and dsHSV24 from (B) were replotted to demonstrate their noncooperative binding profiles. (D) The stoichiometry between IFI16 and dsVACV72 was determined using the concentration of FAM-dsVACV72 sixfold higher than its determined K_{Dapp} (Table 2.2). The lines are fits to a quadratic binding formula. Shown is representative of n = 3. (E) Stoichiometry of IFI16 to dsDNA was determined by plotting the number of bound IFI16 molecules vs. dsDNA length (dsVACV72, dsHSV60, dsVACV52, and dsHSV33) (n = 3). (F) A plot of binding-site normalized K_{Dapp} vs. dsDNA length for IFI16 (n = 3).

DNA Label	DNA Sequence
VACV72	ATCCATCAGAAAGAGGTTTAATATTTTTGTGAGACCATCGAAGA GAGAAAGAGATAAAACTTTTTTACGACT
HSV60	TAAGACACGATGCGATAAAATCTGTTTGTAAAATTTATTAAGGG TACAAATTGCCCTAGC
VACV52	ATCCATCAGAAAGAGGTTTAATATTTTTGTGAGACCATCGAAGA GAGAAAGA
HSV39	TAAGACACGATGCGATAAAATCTGTTTGTAAAATTTATT
HSV33	TAAGACACGATGCGATAAAATCTGTTTGTAAAA
HSV24	TAAGACACGATGCGATAAAATCTG
HSV15	TAAGACACGATGCGA
AG60	AG x 30
dsDNAxxx	xxx-bp PCR fragments of dsDNA

Table 2.1: DNA sequences used in study.

DNA	K_{Dapp}, nM	Normalized K_{Dapp}, nM	Hill Coefficient
dsVACV72	65 ± 19	325	1.9 ± 0.2
dsHSV60	128 ± 31	512	1.5 ± 0.1
dsVACV52	172 ± 32	596	1.3 ± 0.1
dsHSV39	245 ± 21	637	1.2 ± 0.1
dsHSV33	461 ± 53	1,014	1.1 ± 0.1
dsHSV24	864 ± 98	1,382	1.1 ± 0.1
dsHSV15	NB	ND	ND

Table 2.2: Binding of IFI16 against various dsDNA as measured by fluorescence anisotropy. The apparent binding affinities (K_{Dapp} values) and Hill coefficients are mean values of at least three independent FA experiments. Each FAM-labeled DNA was used at least 20-fold lower than the determined K_{Dapp} (n = 3; ± SD). NB, no significant binding was observed up to 2μM protein; ND, not determined.

2.2.2 IFI16 Oligomers Are Distinct Protein Clusters

In contrast to the previous model in which IFI16·dsDNA complexes resemble noninteracting beads on a string [9], these binding experiments suggested that IFI16 assembles into distinct oligomeric clusters on dsDNA. These two models can be distinguished by monitoring formation of FAM-dsDNA·IFI16 complexes in the presence of increasing amounts of unlabeled DNA using native gel electrophoretic mobility-shift assays (EMSAs): the clustering mechanism would show a concerted transition without resulting in significant intermediates, whereas the noninteracting mechanism would clearly display intermediate species [25] [26]. Despite extensive efforts to optimize conditions, IFI16·dsVACV72/HSV60 complexes did not fully enter into the EMSA gel matrix. This was likely caused by the unusually high pI of IFI16 (pI: 9.3), as similar results were observed for AIM2 (pI: 9.8) and other DNA-binding proteins with high pI values [10] [27]; thus analyses of EMSA is limited to qualitative purposes only. Nevertheless, judging by the disappearance of free FAM-dsDNA and by the formation of only a few transient intermediates, these results are consistent with the FA assays in which IFI16 binds dsDNA cooperatively and eventually form an oligomer on each dsDNA in a length-dependent manner (**Figure 2.3A**). I then performed a competition EMSA in which increasing concentrations of unlabeled dsAG60 were added to FAM-dsVACV72·IFI16 (**Figure 2.3B**). As predicted from the clustering model, IFI16 bound dsVACV72 in a concerted manner without forming distinct intermediates. Importantly, the lack of intermediates in this assay indicates that IFI16 preferentially clusters even if there are nonadjacent excess binding sites (i.e., excess dsDNA).

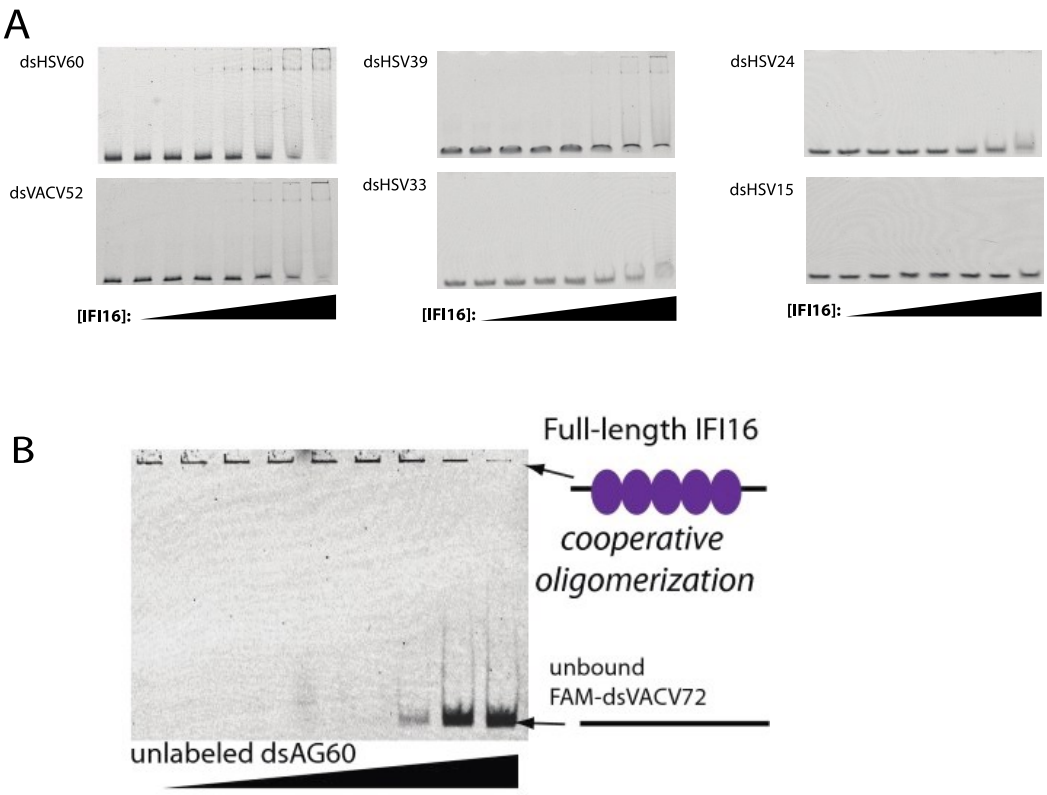

Figure 2.3: IFI16 cooperatively clusters on dsDNA. (A) EMSAs of various dsDNA (5 nM) against increasing concentrations of IFI16 (0, 1.5, 3, 6, 12, 24, 48, and 96 nM). (B)and EMSA in which increasing concentrations of dsAG60 (1.5, 3, 6, 12, 23, 45, 95, and 190 μg/mL) were added to IFI16•FAM-dsVACV72 (190 nM and 0.2 μg/mL respectively).

2.2.3 IFI16 Oligomerizes on dsDNA in a Switch-Like Manner

To quantitatively determine the oligomerization efficiency of IFI16 in the presence of excess DNA, I developed a Förster resonance energy transfer (FRET) assay by labeling one batch of IFI16 with a FRET donor and another with an acceptor. Importantly, because I am adding increasing amounts of excess DNA to a fixed concentration of IFI16, FRET signals are expected to arise only if IFI16 preferentially binds next to one another even in the presence of excess nonadjacent binding sites. Indeed, adding increasing amounts of unlabeled dsVACV72 to an equimolar (20nM each) mixture of FRET-labeled IFI16 produced changes of the emission ratios between the donor (decrease) and acceptor (increase) that could fit to a Hill equation (I denote the midpoints of these curves as K_{Fapp} values, apparent oligomerization constants; **Figure 2.4A,B,C** and **Table 2.3**). No significant FRET signals were observed from ssDNA (e.g., ssVACV72) or dsDNA shorter than 60 bp (e.g., **Figure 2.4A**, inset). The amplitudes of the FRET ratios from dsDNA fragments longer than 72 bp were greater but otherwise did not differ significantly from one another, indicating that all labeled IFI16 molecules used in the assay were bound to these fragments to form similarly sized oligomers (**Figure 2.4B**). After determining the K_{Fapp} value for each dsDNA fragment, I plotted relative binding efficiency (normalized to the K_{Fapp} of dsDNA2000) vs. dsDNA length to analyze how the oligomerization (binding) efficiency of IFI16 changes with respect to the length of dsDNA. This plot shows a cooperative relationship best fit with a Hill coefficient of 5.2 ± 0.6 and the optimal oligomerization efficiency (infliction point) around 150 bp (indicated by the arrow in **Figure 2.4C**). The Hill coefficient of about 5 in this plot

indicates that a small fixed amount of IFI16 can dramatically amplify its oligomerization efficiency with increasing length of dsDNA in a switch-like manner. Moreover, the infliction point at 150 bp suggests that about ten IFI16 molecules comprise an optimal binding cluster (15 bp per one IFI16).

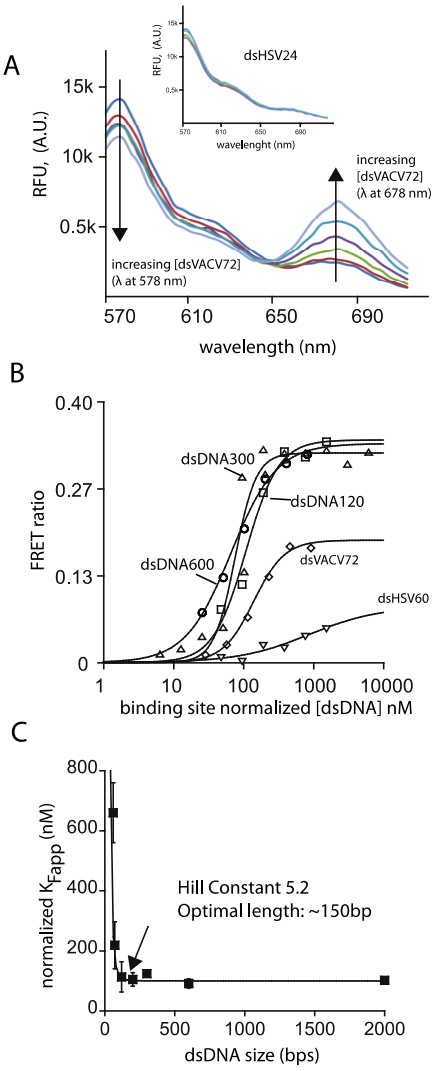

Figure 2.4: IFI16 oligomerizes on dsDNA with a switch-like mechanism. (A) A sample fluorescence emission spectra of an equimolar mixture of FRET donor (Dylight 550)- and acceptor (Dylight 650)-labeled IFI16 with increasing concentrations of dsVACV72 (0, 1.5, 3, 6, 12 24, 48 μg/mL; excitation at 522 nm). A.U., arbitrary units; RFU, relative fluorescence units. Inset: the emission spectra of FRET-labeled IFI16 with increasing concentrations of dsHSV24 (0, 10, 20, 50, 100 μg/mL). (B) A sample plot of changes in the ratio between the FRET donor emission (λ_{max}: 578 nm) and the acceptor emission (λ_{max}: 678 nm) at each indicated dsDNA concentration. The apparent oligomerization constants (K_{Fapp}) were obtained by fitting the data to a Hill equation (Table 2.2), and shown is a representative of n = 3. (C) A plot of binding efficiency vs. the length of dsDNA for IFI16. The data were fit to a Hill equation. The efficiency was determined by normalizing the mean K_{Fapp} of each fragment to that of dsDNA2000. The arrow indicates the infliction point (~150 bp).

DNA	K_{Fapp}, nM	Normalized K_{Fapp}, nM	Hill Coefficient
dsHSV60	165 ± 25	660	1.3 ± 0.2
dsVACV72	46 ± 16	218	1.8 ± 0.1
dsDNA120	14 ± 6	113	1.6 ± 0.2
dsDNA200	7.8 ± 1.7	104	1.8 ± 0.3
dsDNA300	6.2 ± 0.5	124	2.3 ± 0.5
dsDNA600	2.3 ± 0.4	91	1.8 ± 0.1
dsDNA2000	0.8 ± 0.1	102	1.6 ± 0.2

Table 2.3: Oligomerization of IFI16 by various dsDNA as measured by FRET.

2.2.4 HIN200 Domains Bind dsDNA Noncooperatively and Result in Weaker Affinity

To identify the functional domain of IFI16 that drives filament formation, I generated the individual and tandem HIN200 domains of IFI16 ($IFI16^{HinA}$, $IFI16^{HinB}$, and $IFI16^{HinAB}$; **Figure 2.1**) and assayed their binding to FAM-dsVACV72. Surprisingly, I did not observe significant dsDNA binding from these IFI16 variants (e.g., **Figure 2.5A**, *Upper*). It was previously reported that dsDNA-binding affinity of $IFI16^{HinB}$ was significantly influenced by buffer salt concentrations [9]. Indeed, I found that the near-physiological salt concentrations (160 mM KCl) in the reaction buffer interfered with binding (**Figure 2.5A**, *Lower*). Unlike full-length, the disappearance of free FAM-dsVACV72 and dsHSV60 was apparently noncooperative even under 60 mM KCl, suggesting that the DNA-bound complexes formed by the tandem HIN200 domains of IFI16 must resemble noninteracting beads on a string (**Figure 2.5A,B**). To detect binding of the individual HIN200 domains, I decreased the KCl concentration even further to 20 mM. Here, I found that all IFI16 variants but $IFI16^{PYD}$ bound dsDNA in decreasing order of affinity: full-length, $IF16^{HinAB}$, $IFI16^{HinA}$, $IFI16^{HinB}$. Importantly, as expected from the EMSA, the fits to the isolated HIN200 domains binding to dsVACV72 were noncooperative (Hill coefficients, ~1, **Figure 2.5C**). These results suggest the HIN200 domains of IFI16, individually or in tandem, do not efficiently bind dsDNA because they fail to cooperatively oligomerize.

Several additional experiments performed at 60 mM KCl confirmed that the tandem HIN200 domains of IFI16 do not oligomerize. First, in FA binding assays, I

found that unlike full-length, IFI16^{HinAB} bound all FAM-dsDNA fragments noncooperatively and without apparent length-dependency (**Figure 2.5D,E** vs. **Figure 2.2**). Second, in an EMSA competition experiment similar to **Figure 2.3B**, IFI16^{HinAB} clearly displayed intermediates (**Figure 2.5F**). Third, no significant FRET signals were observed from donor- and acceptor-labeled IFI16^{HinAB} using increasing concentrations of dsVACV72 (**Figure 2.5G** vs. **Figure 2.4A**). These results consistently support the idea that the HIN200 domains of IFI16 do not cooperatively cluster on dsDNA but independently bind dsDNA analogous to beads on a string.

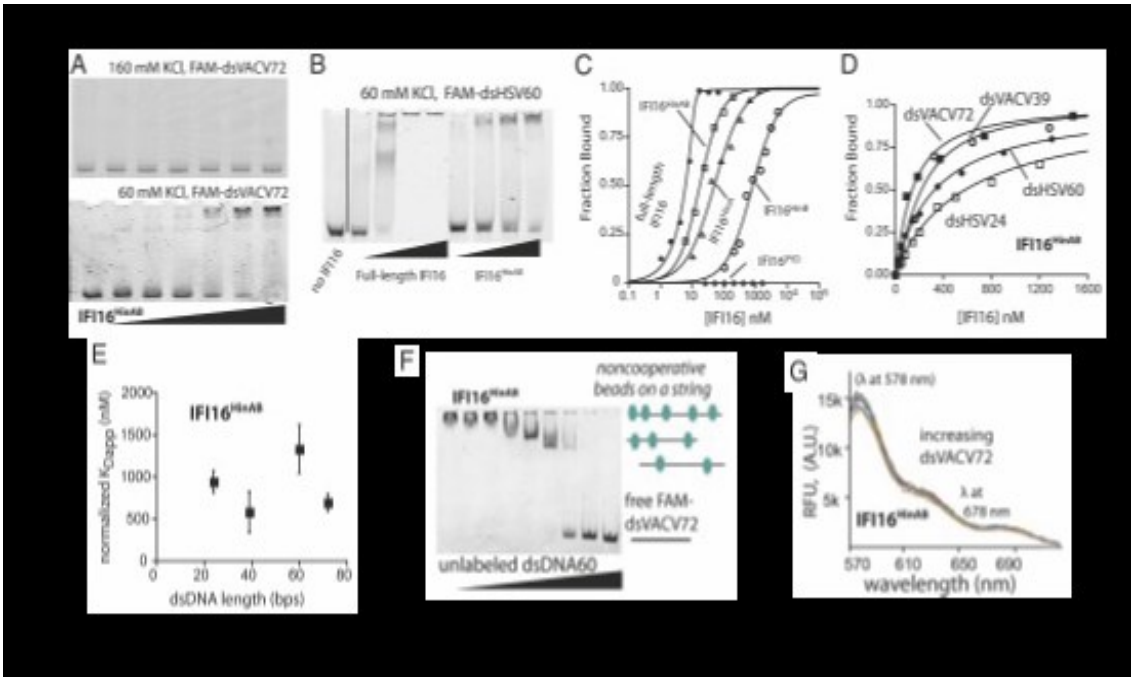

Figure 2.5: The HIN200 domains of IFI16 do not oligomerize on dsDNA. (A) An EMSA of IFI16^{HinAB} to FAM-dsVACV72 (2 nM) at 160 mM KCl (*Upper*) and 60 mM KCl (*Lower*) (IFI16^{HinAB}: 0, 1.5, 3, 6, 12, 24, 48, and 96 nM). (B) An EMSA of full-length IFI16 (50, 100, 200, 400 nM) and IFI16^{HinAB} (100, 200, 400, 800 nM) to FAM-dsHSV60 (20 nM). (C) Binding of IFI16 variants to FAM-dsVACV72 were determined by FA at 20 mM KCl. The determined values are listed in Table 2.3. The data for full-length IFI16 were fit with a quadratic form of binding equation, and the others were fit with a Hill equation. Shown is a representative of n = 4. (D) Binding of IFI16^{HinAB} toward indicated FAM-dsDNA fragments (5 nM each) was determined by FA at 60 mM KCl. The lines are fits to a Hill equation. Shown is a representative of four independent experiments. (E) Binding-site normalized K_{Dapp} values of IFI16^{HinAB} from *D* (n = 4; ±SD). (F) An EMSA in which increasing concentrations of dsAG60 (190, 95, 45, 23, 12, 6, 3, and 1,5 μg/mL) were added to IFI16^{HinAB}·FAM-dsVACV72 (250 nM and 0.2 μg/mL, respectively) using 60 mM KCl. (G) The lack of FRET signals for donor- and acceptor-labeled IFI16^{HinAB} (100 nM each) with increasing dsVACV72 (0, 0.5, 1, 2, 4, 8, 17, 35, 70, 140 μg/mL) at 60 mM KCl.

2.2.5 The PYD of IFI16 is Necessary for Cooperative DNA Binding

Conventionally, the role of PYDs in ALRs is thought to be limited to recruiting downstream effectors [1] [19] [28]. However, several independent experiments I have performed thus far suggested that IFI16^{PYD} is important for oligomerization on dsDNA. To test this possibility, I first confirmed that IFI16^{PYD} is a monomer and does not directly bind dsDNA (**Figure 2.6A**; see also **Figure 2.1A**). Next, I constructed IFI16 variants consisting of either IFI16^{PYD} and IFI16^{HinA} (IFI16$^{\Delta B}$; residues 1–393) or IFI16^{PYD} and IFI16^{HinB} (IFI16$^{\Delta A}$; residues 1–191 plus 459–729) (see also **Figure 2.1**). At 160 mM KCl, I found that IFI16$^{\Delta B}$ efficiently binds dsVACV72 and IFI16$^{\Delta A}$ binds dsVACV72 more weakly than IFI16$^{\Delta B}$, likely reflecting the weaker binding affinity of the IFI16^{HinB} in my hands compared with the previous report [3] (**Figure 2.6A,B** and **Table 2.4**). Importantly, unlike the IFI16 constructs without the PYD, the binding profiles were positively cooperative (**Figure 2.6B** and **Table 2.4**).

PYDs are homotypic protein-protein interaction domains in which multiple binding surfaces are used in their interactions [29]. For example, helices one and four dock on helices two and three to create a high-affinity oligomeric interaction [30] (see also **Figure 2.6D**). To identify surface residues in IFI16^{PYD} important for the cooperative filament formation, I generated a set of mutations on the conserved surface residues using a homology model of IFI16^{PYD} based on the PYD of myeloid cell nuclear differentiation antigen (MNDA) (Protein Data Bank ID code 2DBG; 88% sequence similarity; **Figure 2.6C,D**). I found that mutating these surface side-chains on IFI16^{PYD} resulted in up to 25-fold increase in K_{Dapp} values for FAM-dsVACV72 compared with wild-type (**Figure 2.6E** and **Table 2.4**). In addition, the binding

profiles of IFI16^{mut2} and IFI16^{mut3} were essentially noncooperative (Hill coefficients: ~1 for both; **Figure 2.6E**), indicating that the cooperative dsDNA-binding mechanism requires intact IFI16^{PYD}. Also supporting the FA results, IFI16^{mut3} bound FAM-dsVACV72 significantly weakly compared with wild-type with an altered migration pattern in an EMSA (**Figure 2.6F**, *Upper*). Moreover, unlike wild-type, IFI16^{mut3} showed intermediates in a competition ESMA (**Figure 2.6F**, *Lower*). Finally, no FRET signal was observed between an equal mixture of donor-labeled IFI16^{HinAB} and acceptor-labeled full-length IFI16 with increasing dsVACV72 at 60 mM KCl (**Figure 2.6G**). The lack of FRET signal in this experiment suggests that IFI16^{PYD} interacts with one another instead of interacting with adjacent HIN200 domains on the dsDNA scaffold. Taken together, I concluded that the conditional proximity induced upon encountering dsDNA by the HIN200 domains triggers IFI16^{PYD}-driven cooperative filament formation of IFI16.

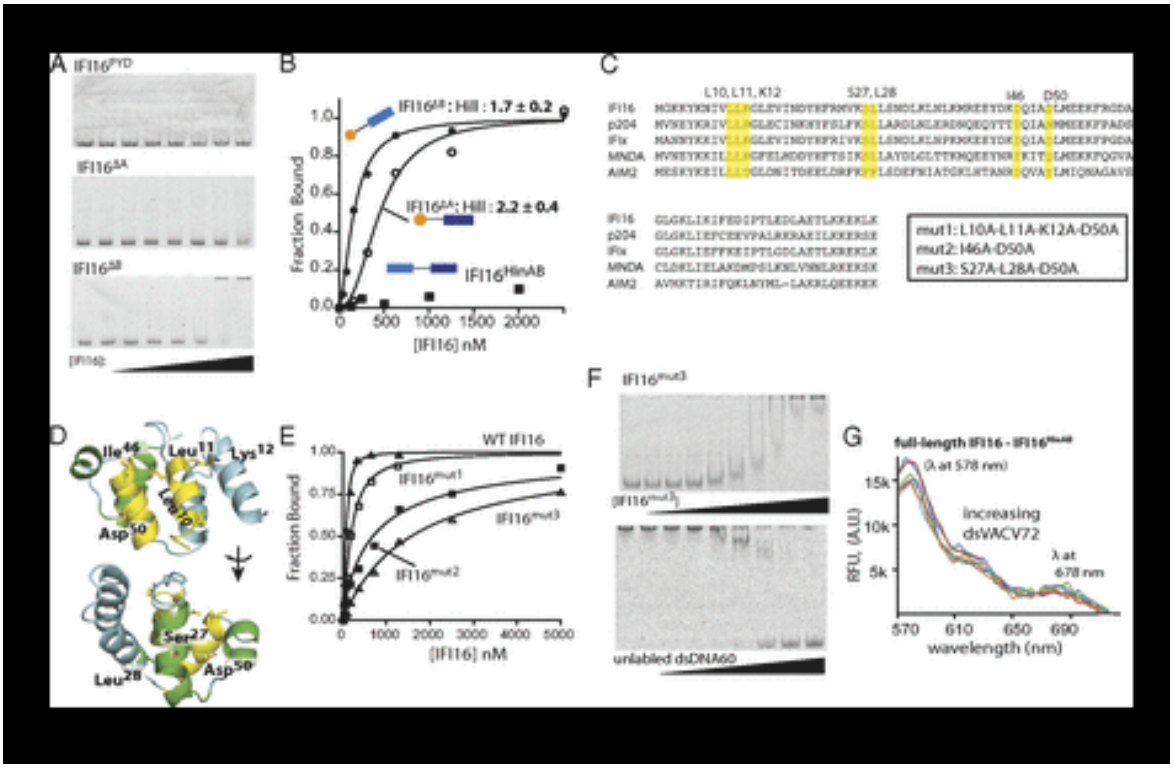

Figure 2.6: IFI16^{PYD} plays a positive role in cooperative oligomerization. (A) EMSAs of IFI16^{PYD} (*Top*), IFI16$^{\Delta A}$ (*Middle*), and IFI16$^{\Delta B}$ (*Bottom*) to FAM-dsVACV72 at 160 mM KCl (0, 1.5, 3, 6, 12, 24, 48, and 96 nM; for IFI16^{PYD}: 15, 30, 62, 125, 250, 500, and 1,000 nM). (B) Binding of IFI16^{HinAB}, IFI16$^{\Delta A}$, and IFI16$^{\Delta B}$ to FAM-dsVACV72 was determined by FA at 160 mM KCl. The data were fit with a Hill equation (Table 2.4). Shown is a representative of n = 3. (C) A sequence alignment of the PYDs of various ALRs. The mutated residues are highlighted, and the alignment was generated using ClustalOmega (www.clustal.org). P204 is a mouse homolog of IFI16, and all of the other proteins are human ALRs. (D) A homology model of IFI16^{PYD} was generated using the PYD of MNDA as a template and the SWISS-MODEL server. Mutated side chains are shown with a stick representation; the figure was generated using PyMOL (DeLano Scientific). (E) Binding of each mutant IFI16 to FAM-dsVACV72 was determined by FA at 160 mM KCl. The wild-type data from Figure 2A are shown for comparison. The determined values are listed in Table 2.4, and shown is a representative of n = 3. (F, *Upper*) an EMSA-binding assay of IFI16^{mut3} (0, 3, 6, 12, 24, 48, 96, 192, 384, 768 nM) to FAM-dsVACV72 (0.4 μg/mL). (F, *Lower*) An EMSA in which increasing concentrations of dsAG60 (0, 1.5, 3, 6, 12, 23, 45, and 190 μg/mL) were added to IFI16^{mut3}·FAM-dsVACV72 (768 nM and 0.2 μg/mL, respectively). (G) The lack of FRET signals from an equal mixture of donor-labeled IFI16^{HinAB} and acceptor-labeled full-length IFI16 (50 nM each) with increasing dsVACV72 (0, 1.2, 6, 28, 140 μg/mL) at 60 mM KCl.

IFI16 Construct	K_{Dapp}, nM	Hill Coefficient
20mM KCl		
Full-Length	0.5*	ND
$IFI16^{HinA-B}$	12 ± 2	1.0 ± 0.1
$IFI16^{HinA}$	51 ± 8	1.0 ± 0.1
$IFI16^{HinB}$	768 ± 87	1.0 ± 0.1
$IFI16^{PYD}$	NB	ND
160mM KCl		
Full-Length	65 ± 19	1.9 ± 0.2
$IFI16^{\Delta A}$	458 ± 83	2.2 ± 0.4
$IFI16^{\Delta B}$	197 ± 45	1.7 ± 0.2
$IFI16^{D50A}$	77 ± 15	1.8 ± 0.2
$IFI16^{mut1}$	195 ± 24	1.5 ± 0.1
$IFI16^{mut2}$	708 ± 87	0.8 ± 0.2
$IFI16^{mut3}$	1,580 ± 220	1.0 ± 0.1

Table 2.4: Binding of IFI16 constructs to dsVACV72

2.3 Discussion

2.3.1 What Could Be the Molecular Basis for DNA Length-Dependent Responses?

Conventionally, it was thought that IFI16 and its related ALRs prefer long dsDNA as it can accommodate a greater number of ALRs, which would be required to promote multimeric protein·protein interactions with their downstream partners [3] [9] [18]. However, an important point to consider here is that if IFI16 binds dsDNA resembling beads on a string as previously thought, longer DNA fragments can be as counterproductive for multimeric interactions as short fragments. For example, even one femtogram of transfected 100-bp DNA fragments (≥9,000 copy number) provide more than 100,000 equally plausible nonadjacent binding sites for the HIN200 domains of IFI16. Furthermore, in principle, one HSV-1 genome contains more than 30,000 binding sites. Thus, because IFI16 lacks any DNA sequence specificity, without an intrinsic clustering mechanism mediated by protein·protein interactions, IFI16 would be randomly scattered on any self- or nonself-DNA (**Figure 2.7A**) instead of selectively forming signaling foci on foreign DNA as observed *in vivo* [11] [12] [13] [14]. These results suggest that the PYD-driven filament formation by IFI16 can provide an effective mechanism that allows multiple IFI16 (or its related ALRs) to bind adjacent to one another to form signaling foci even in the presence of excess DNA (**Figure 2.7B**).

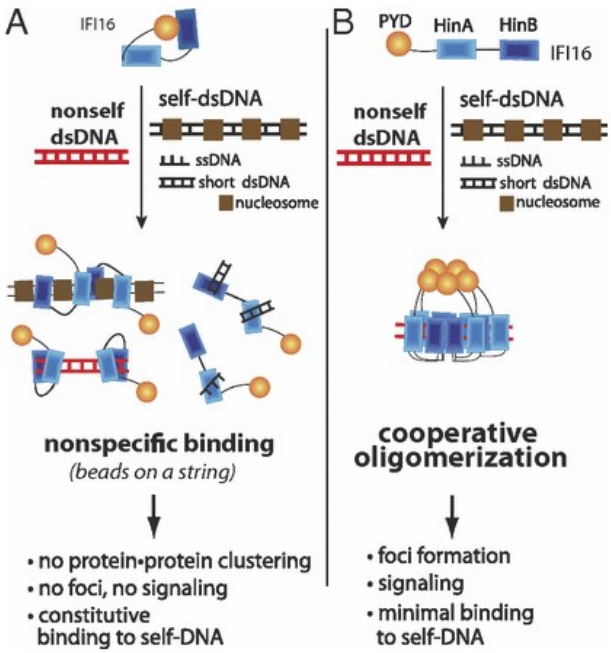

Figure 2.7: Two models for IFI16. (A) The conventional model lacks any regulatory mechanism that can account for the observed behavior of IFI16 *in vivo*. (B) A competing model established based on the results presented in this study.

2.3.2 What Is a Potential Regulatory Mechanism of IFI16 in the Nucleus?

A striking behavior of IFI16 is that even though it is localized in the nucleus, it is randomly diffused without forming distinct foci with self-DNA [11] [12] [13] [14]. In fact, one of the longest-standing questions regarding IFI16 is how it does not engage self-DNA in the nucleus [11] [12] [13] [14] [31] [32]. I show here that the HIN200 domains of IFI16 possess negligible DNA-binding capacities under physiologically relevant salt concentrations [i.e., K_D >> 10 μM for the HIN domains in the ~200 mM effective salt concentrations of the host cell nucleus [33]]. Importantly, the interaction between the HIN200 domains and dsDNA are exclusively mediated by the phosphate backbone [i.e., nonspecific electrostatics [9]]. Although useful in engaging a wide variety of foreign DNA, this intrinsically weak intermolecular force does not generate enough binding energy beyond formation of an encounter complex [34]. Thus, despite exceeding the footprint of the HIN200 domains, the length of the exposed linker-dsDNA between nucleosomes [10–20 bp [21]] or even that of the transcription bubble [~17 bases [22]] is too short to promote robust filament assembly of IFI16. Additionally, I envision that requiring oligomerization to achieve tight binding also plays a negative role in competing against replication/transcription machinery. Thus, these data suggest that the weak binding capacity of its HIN200 domains coupled with filament formation could sufficiently suppress its interaction with self-DNA.

2.3.3 How Can IFI16 Selectively Engage Foreign DNA and Assemble into Large Signaling Foci?

IFI16 is known to selectively colocalize with viral genomic DNA to form signaling foci within the nucleus (e.g., HSV-1, KSHV) [11] [12] [13] [14]. Although the total amount of viral DNA would never exceed that of self-DNA, there are two features of foreign DNA that are much more conducive to filament formation by IFI16. First, the entire naked DNA genome is exposed immediately after the invasion [35] [36] [37]. Second, although the viral genome packages into chromatin with the host histones, it is less dense and much more loosely packed than the nuclear self-DNA [35] [36] [37]. Importantly, I find that the relative binding affinity of IFI16 to various dsDNA fragments changes cooperatively not only with increasing IFI16 concentrations (**Figure 2.2**) but also with increasing number of binding sites (dsDNA lengths; **Figure 2.4**). These highly cooperative relationships suggest that IFI16 is capable of amplifying its clustering behavior in a switch-like manner. For example, the Hill coefficient of about 5 from **Figure 2.4C** suggests that even the low basal amount of IFI16 [38] prefers by more than 2,000-fold to oligomerize on a 150-bp fragment rather than simply binding to a 15-bp fragment. Additionally, by these cooperative mechanisms, the binding efficiency diminishes in the same manner it is amplified; thus, these results suggest that IFI16 can clearly define an “on” or “off” state with respect to its concentration and the length of dsDNA. Collectively, the filament formation by IFI16 provides a compelling mechanism by which it could selectively engage foreign DNA while minimizing its interaction with self-DNA.

2.3.4 The Role of PYD.

In contrast to the autoinhibitory role of the PYD of AIM2 [9] [18], I find that IFI16^{PYD} plays an unexpected positive role in dsDNA binding. Several surface residues in IFI16^{PYD} that mediate the cooperative dsDNA binding of IFI16 are highly conserved (**Figure 2.6C**), thus suggesting that other related ALRs use a common oligomerization mechanism. Interestingly, I find that the PYD-driven filament assembly could provide two important tactical advantages that might not be attainable by the previously proposed noninteracting model [9]. First, it allows formation of an ordered array of IFI16^{PYD} oligomers on foreign DNA even with the low basal concentration of IFI16 *in vivo* [38]. The IFI16^{PYD} cluster would then be instrumental for subsequent steps in IFI16 signaling pathways, because assembling PYD/CARD oligomers by receptor molecules is required for recruiting downstream partners [23] [24] [39]. Second, the PYD-driven filament assembly ensures formation of a high-affinity IFI16·foreign DNA complex for signaling, which would also directly interfere with replication of pathogens. Indeed, IFI16 can directly suppress replication of human cytomegalovirus (HCMV) without downstream activation of IFN-β [40]. The authors of this report [40] also showed that this intrinsic antiviral activity of IFI16 required its PYD, supporting the positive role in engaging foreign DNA *in vivo*. Finally, further consistent with the positive role of IFI16^{PYD}, very recent results published while this paper was under review showed that IFI16 oligomerizes on HCMV genome via its PYD in human fibroblasts and that a viral protein (pUL83) sequesters IFI16^{PYD} to evade the host immune response [41]. Future

studies could reveal how PYDs regulate the function of IFI16 and its related ALRs in more detail.

2.3.5 Filament Formation as a Broad Host Defense Strategy.

Conventionally, assembling filaments on foreign nucleic acids has been considered as a defense strategy reserved for sensing intracellular RNA [28] [42]. For example, a principal mechanism by which melanoma differentiation associated protein (MDA)5 distinguishes self- from nonself-RNA is via cooperatively assembling into filaments along the length of long dsRNA with its two CARDs forming clusters tracing the center of filaments in a helical trajectory [25] [26] [42] [43] [44] [45]. IFI16 and MDA5 are unrelated proteins, and, not surprisingly, mechanisms allowing these proteins to assemble into filaments are significantly different. For example, the RecA-like RNA-sensing domain of MDA5 wraps around dsRNA and promotes filament formation in a head-to-tail manner; the CARD clusters are not required for filament assembly. In contrast, the DNA-binding domains of IFI16 do not directly promote filament formation, and yet its PYD is necessary to drive this process. Interestingly, the data suggest that an optimal binding unit consists of approximately ten IFI16 protomers (**Figure 2.4C**). It is tempting to speculate that the filament assembly is then accomplished by propagating these decameric units along the length of dsDNA. Another difference between IFI16 and MDA5 is that the helicase domain regulates the lifetime of MDA5 filaments via ATP hydrolysis, although it is not clear what triggers the disassembly of IFI16 filaments. Further structural and kinetic studies could reveal the architecture and dynamics of IFI16 filaments in more detail. Overall, regardless of these mechanistic differences, the

results suggest that assembling filaments of intracellular dsDNA/dsRNA receptors is not only an effective mechanism but also a broadly conserved host defense strategy against foreign nucleic acids.

2.4. Methods

2.4.1 Protein Expression and Purification.

Human IFI16 full-length (residues 1–729), IFI16^{PYD} (residues 1–93), IFI16^{HinA} (residues 192–393), IFI16^{HinB} (residues 518–729), IFI16^{HinAB} (residues 192–729), IFI16$^{\Delta A}$ (residues 1–191 + 459–729), and IFI16$^{\Delta B}$ (residues 1–393) were cloned into a pET21b vector (Novagen) using standard PCR methods. The identity of each IFI16 construct and mutant was confirmed by DNA sequencing. IFI16 constructs were expressed using T7 express cells (NEB). The cells were grown until OD_{600}: 0.5 at 37 °C and protein expression was induced by 0.2 mM isopropyl β-d-1-thiogalactopyranoside overnight at 18 °C. Each protein construct was purified using Ni^{2+}-nitrilotriacetic acid (NTA) affinity chromatography (Qiagen), followed by cation exchange chromatography using a Resource S column (GE Healthcare) with a NaCl gradient (50 mM to 1 M) and Superdex-200 or Superdex-75 size-exclusion chromatography (GE Healthcare). Micrococcal nuclease (NEB; 100 units per 1 L of cells) was added during cell lysis to eliminate bacterial DNA contamination. All proteins were eluted as monomers in size-exclusion chromatography (**Figure 2.1A**), and UV-visible spectroscopy confirmed that they were free from nucleic acid contamination (**Figure 2.1B**). Each protein was greater than 95% pure (**Figure 2.1C**) and concentrated and stored at −80 °C in 20 mM Hepes·NaOH at pH 7.4, 200 mM NaCl, 1 mM EDTA, 5 mM DTT, and 10% (vol/vol) glycerol.

2.4.2 DNA.

Each FAM-labeled and unlabeled DNA fragment shorter than 120 bp was commercially purchased from Integrated DNA Technologies (IDT; **Table 2.1**).

Duplex formation was achieved by mixing the sense and the complementary strands in 1:1 molar ratio in buffer A (see below) and heating in 95 °C for 10 min. The duplex samples were subsequently cooled to room temperature (25 ± 2 °C) on the bench top. The longer dsDNA fragments were generated by PCR using a plasmid. Under the reaction conditions, all duplex DNAs are at least 15 °C below their predicted melting temperatures (SciTools; IDT).

2.4.3 Biochemical Assays.

Unless noted otherwise, all assays were performed at room temperature (25 ± 2 °C) with 40 mM Hepes·NaOH at pH 7.4, 160 mM KCl, 1 mM EDTA, 2 mM DTT, 10% glycerol, and 0.1% Triton-X100 (referred to as buffer A). The FA of FAM-labeled DNA (at least 15- to 20-fold lower concentrations than the K_{Dapp} for each experiment) was monitored with increasing concentrations of IFI16 variants by using a Tecan M1000 plate reader (excitation at 470 nm and emission at 528 nm) and 384-well nonbinding plates (Corning), and FA values were determined by iControl data analysis software (Tecan). The concentration of each FAM-dsDNA was kept at least 20-fold less than its respective K_{Dapp} in each binding assay to prevent any length bias (typically 2.5–5 nM per assay). For EMSA, each DNA was incubated with indicated concentrations of IF16 variants for 10 min and the bound complexes were separated from unbound dsDNA using 6% Tris·boric acid-EDTA (TBE) gels with 0.25× TBE as a running buffer. The results were then visualized by scanning on a Typhoon imager (GE Healthcare; excitation at 488 nm; emission at 532 nm). FRET assays were also performed in a 384-well plate format using the Tecan M1000. All experiments were performed at least three times independent of one another, and

errors were calculated by using the SDs (number of independent experiments, $n \geq 3$). A Hill, quadratic, or competition form of binding equation were fit to the data using Kaleidagraph (Synergy Soft).

2.4.4 Protein Labeling.

IFI16 and IFI16^{HinAB} were labeled using twofold molar excess of Dylight 550 maleimide or Dylight 650 maleimide (Pierce) in buffer A with DTT replaced by TCEP. Labeling reactions were quenched by adding excess β-mercaptoethanol and the labeled-proteins were purified using Ni^{2+}-NTA chromatography followed by size-exclusion chromatography. The labeling efficiency of IFI16 to each dye was approximately one to one, and there was no considerable difference in FRET data from two different batches of labeled proteins.

2.5 References

[1] SA Schattgen and KA Fitzgerald, "The PYHIN protein family as mediators of host defenses," *Immunology Reviews*, vol. 243, no. 1, pp. 109-118, 2011.

[2] SR Paludan and AG Bowie, "Immune sensing of DNA," *Immunity*, vol. 38, no. 5, pp. 870-880, 2013.

[3] L Unterholzner et al., "IFI16 is an innate immune sensor for intracellular DNA," *Nat Immunol*, vol. 11, no. 11, pp. 997-1004, Nov 2010.

[4] S Veeranki and D Choubey, "Interferon-inducible p200-family protein IFI16, an innate immune sensor for cytosolic and nuclear double-stranded DNA: Regulation of subcellular localization," *Mol Immunol.*, vol. 49, no. 4, pp. 567-571, 2012.

[5] M Mondini et al., "The interferon-inducible HIN-200 gene family in apoptosis and inflammation: implication for autoimmunity," *Autoimmunity*, vol. 43, no. 3, pp. 226-31, Apr 2010.

[6] K Uchida et al., "Identification of specific autoantigens in Sjögren's syndrome by SEREX," *Immunology*, vol. 116, no. 1, pp. 53-63, Sept 2005.

[7] M Mondini et al., "Role of the interferon-inducible gene IFI16 in the etiopathogenesis of systemic autoimmune disorders," *Ann N Y Acad Sci*, vol. Sept, no. 1110, pp. 47-56, 2007.

[8] F Gugliesi et al., "Nuclear DNA sensor IFI16 as circulating protein in autoimmune diseases is a signal of damage that impairs endothelial cells through high-affinity membrane binding," *PLoS One*, vol. 8, no. 5, p. e63045, May 2013.

[9] T Jin et al., "Structures of the HIN domain:DNA complexes reveal ligand binding and activation mechanisms of the AIM2 inflammasome and IFI16 receptor," *Immunity*, vol. 36, no. 4, pp. 561-71, Apr 2012.

[10] T Fernandes-Alnemri, JW Yu, P Datta, J Wu, and ES Alnemri, "AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA," *Nature*, vol. 458, no. 7237, pp. 509-13, Mar 2009.

[11] N Kerur et al., "IFI16 acts as a nuclear pathogen sensor to induce the inflammasome in response to Kaposi Sarcoma-associated herpesvirus infection," *Cell Host Microbe*, vol. 9, no. 5, pp. 363-75, May 2011.

[12] T Li, BA Diner, J Chen, and IM Cristea, "Acetylation modulates cellular distribution and DNA sensing ability of interferon-inducible protein IFI16," *Proc Natl Acad Sci U S A*, vol. 109, no. 26, pp. 10558-63, Jun 2012.

[13] MH Orzalli, NA DeLuca, and DM Knipe, "Nuclear IFI16 induction of IRF-3 signaling during herpesviral infection and degradation of IFI16 by the viral ICP0 protein," *Proc Natl Acad Sci U S A*, vol. 109, no. 44, pp. E3008-17, Oct 2012.

[14] VV Singh et al., "Kaposi's sarcoma-associated herpesvirus latency in endothelial and B cells activates gamma interferon-inducible protein 16-mediated inflammasomes," *J Virol*, vol. 87, no. 8, pp. 4417-31, Apr 2013.

[15] H Yan et al., "RPA nucleic acid-binding properties of IFI16-HIN200," *Biochim Biophys Acta*, vol. 1784, no. 7-8, pp. 1087-97, Jul-Aug 2008.

[16] V Brázda, J Coufal, LC Liao, and CH Arrowsmith, "Preferential binding of IFI16 protein to cruciform structure and superhelical DNA," *Biochem Biophys Res Commun*, vol. 422, no. 4, pp. 716-20, Jun 2012.

[17] JW Jones et al., "Absent in melanoma 2 is required for innate immune recognition of Francisella tularensis," *Proc Natl Acad Sci U S A*, vol. 107, no. 21, pp. 9771-6, May 2010.

[18] T Jin, A Perry, P Smith, J Jiang, and TS Xiao, "Structure of the absent in melanoma 2 (AIM2) pyrin domain provides insights into the mechanisms of AIM2 autoinhibition and inflammasome assembly," *J Biol Chem*, vol. 288, no. 19, pp. 13225-35, May 2013.

[19] ES Alnemri, "Sensing cytoplasmic danger signals by the inflammasome," *J Clin Immunol*, vol. 30, no. 4, pp. 512-9, Jul 2010.

[20] JC Liao et al., "Interferon-inducible protein 16: insight into the interaction with tumor suppressor p53," *Structure*, vol. 19, no. 3, pp. 418-29, Mar 2011.

[21] M Kato et al., "Dinucleosome DNA of human K562 cells: experimental and computational characterizations," *J Mol Biol*, vol. 332, no. 1, pp. 111-25, Sept 2003.

[22] M Pal, AS Ponticelli, and DS Lusa, "The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II," *Mol Cell*, vol. 19, no. 1, pp. 101-10, Jul 2005.

[23] H Wu, "Higher-order assemblies in a new paradigm of signal transduction," *Cell*, vol. 153, no. 2, pp. 287-92, Apr 2013.

[24] F Martinon, A Mayor, and J Tschopp, "The inflammasomes: guardians of the body," *Annu Rev Immunol*, vol. 27, pp. 229-65, 2009.

[25] A Peisley et al., "Kinetic mechanism for viral dsRNA length discrimination by MDA5 filaments," *Proc Natl Acad Sci U S A*, vol. 109, no. 49, pp. E3340-9, Dec 2012.

[26] A Peisley et al., "Cooperative assembly and dynamic disassembly of MDA5 filaments for viral dsRNA recognition," *Proc Natl Acad Sci U S A*, vol. 108, no. 52, pp. 21010-5, Dec 2011.

[27] I Leiros, J Timmins, DR Hall, and S McSweeney, "Crystal structure and DNA-binding analysis of RecO from Deinococcus radiodurans," *EMBO J*, vol. 24, no. 5, pp. 906-18, Mar 2005.

[28] E Latz, TS Xiao, and A Stutz, "Activation and regulation of the inflammasomes," *Nat Rev Immunol*, vol. 13, no. 6, pp. 397-411, Jun 2013.

[29] HH Park et al., "The death domain superfamily in intracellular signaling of apoptosis and inflammation," *Annu Rev Immunol*, vol. 25, pp. 561-86, 2007.

[30] PR Vajjhala, RE Mirams, and JM Hill, "Multiple binding sites on the pyrin domain of ASC protein allow self-association and interaction with NLRP3 protein," *J Biol Chem*, vol. 287, no. 50, pp. 41732-43, Dec 2012.

[31] L Unterholzner, "The interferon response to intracellular DNA: why so many receptors?," *Immunobiology*, vol. 218, no. 11, pp. 1312-21, Nov 2013.

[32] L Unterholzner and AG Bowie, "Innate DNA sensing moves to the nucleus," *Cell Host Microbe*, vol. 9, no. 5, pp. 351-3, May 2011.

[33] E Arbely et al., "Acetylation of lysine 120 of p53 endows DNA-binding specificity at effective physiological salt concentration," *Proc Natl Acad Sci U S A*, vol. 108, no. 20, pp. 8251-6, May 2011.

[34] G Schreiber and AR Fersht, "Rapid, electrostatically assisted association of proteins," *Nat Struct Biol*, vol. 3, no. 5, pp. 427-31, May 1996.

[35] AR Cliffe and DM Knipe, "Herpes simplex virus ICP0 promotes both histone removal and acetylation on viral DNA during lytic infection," *J Virol*, vol. 82, no. 24, pp. 12030-8, Dec 2008.

[36] JJ Lacasse and LM Schang, "During lytic infections, herpes simplex virus type 1 DNA is in complexes with the properties of unstable nucleosomes," *J Virol*, vol. 84, no. 4, pp. 1920-33, Feb 2010.

[37] J Oh and NW Fraser, "Temporal association of the herpes simplex virus genome with histone proteins during a lytic infection," *J Virol*, vol. 82, no. 7, pp. 3530-7, Apr 2008.

[38] S Veeranki, X Duan, R Panchanathan, H Liu, and D Choubey, "IFI16 protein mediates the anti-inflammatory actions of the type-I interferons through suppression of activation of caspase-1 by inflammasomes," *PLoS One*, vol. 6, no. 10, p. e27040, 2011.

[39] S Pagliusi et al., "Developing Countries Vaccine Manufacturers Network (DCVMN): engaging to step up for vaccine discovery and access. Meeting report

2012.," *Vaccine*, vol. 31, no. 31, pp. 3111-5, Jun 2013.

[40] GR Gariano et al., "The intracellular DNA sensor IFI16 gene acts as restriction factor for human cytomegalovirus replication," *PLoS Pathog*, vol. 8, no. 1, p. e1002498, Jan 2012.

[41] T Li, J Chen, and IM Cristea, "Human cytomegalovirus tegument protein pUL83 inhibits IFI16-mediated DNA sensing for immune evasion," *Cell Host Microbe*, vol. 14, no. 5, pp. 591-9, Nov 2013.

[42] IC Berke, Y Li, and Y Modis, "Structural basis of innate immune recognition of viral RNA," *Cell Microbiol*, vol. 15, no. 3, pp. 386-94, Mar 2013.

[43] M Wu et al., "Structural basis for dsRNA recognition, filament formation, and antiviral signal activation by MDA5," *Cell*, vol. 152, no. 1-2, pp. 276-89, Jan 2013.

[44] IC Berke, X Yu, Y Modis, and EH Egelman, "MDA5 assembles into a polar helical filament on dsRNA," *Proc Natl Acad Sci U S A*, vol. 109, no. 45, pp. 18437-18441, 2012.

[45] IC Berke and Y Modis, "MDA5 cooperatively forms dimers and ATP-sensitive filaments upon binding double-stranded RNA," *EMBO J*, vol. 31, no. 7, pp. 1714-1726, 2012.

Chapter 3

AIM2 Assembles into Filaments Upon DNA Binding and at High Concentrations: A Mechanism for Activating ASC

Reproduced in part from:

Morrone, SR, Matyszewski, M, Yu, X, Delannoy, M, Egelman, EH, & Sohn, J. (2015). Assembly-driven activation of the AIM2 foreign-dsDNA sensor provides a polymerization template for downstream ASC. *Nature Communications*, 6, 7827.

3.1 Introduction

In the innate immune system of mammals, supramolecular signalling platforms are directly assembled on intracellular foreign double-stranded (ds) DNA and RNA arising from invading pathogens (see refs [1] [2] [3] [4] [5] [6] for review). These supra-structures then induce the sequential polymerization of downstream effectors to propagate upstream signals [7] [8] [9] [10]. Though essential for defense against a number of pathogens such as *Francisella tularensis* and herpes simplex viruses [11] [12] [13] dysregulated foreign-nucleic acid-sensing pathways are associated with several autoimmune disorders including systemic lupus erythematosus and Sjögren's syndrome [14] [15] [16]. How the assembly of these large, complex structures is initiated on appropriate nucleic acids, and how the upstream ligand·receptor assemblies promote the sequential oligomerization of specific downstream effectors are two major unresolved mechanistic questions in understanding the foreign-nucleic acid-sensing pathways [5] [6]. Here, these fundamental questions in the assembly of the foreign-dsDNA-sensing filamentous superstructures by absent-in-melanoma-2 (AIM2) will be explored.

AIM2 is the titular member of the AIM2-like receptor (ALR) family, which also includes other major foreign-dsDNA sensors such as interferon-inducible protein 16 (IFI16) [17] [18] [19] [20] [21]. ALRs directly assemble filamentous signalling platforms (the inflammasomes) on foreign dsDNA [7] [17] [19] [20] [18] [22] [23] [24]. For instance, AIM2 oligomerizes on cytosolic dsDNA and nucleates the polymerization of the adaptor protein ASC (apoptosis-associated speck-forming protein containing a CARD [caspase-recruiting domain]) filament, which then

nucleates the polymerization of the procaspase-1 filament; this final polymerization step activates caspase-1 via auto-proteolysis, triggering inflammatory responses including cytokine maturation and pyroptosis (**Figure 3.1**) [7] [8] [17] [19] [18] [20].

Many proteins build nucleoprotein filaments, frequently by classic RecA-like mechanisms [9] [10] [25] [26]. For instance, retinoic acid-inducible gene-I (Rig-I) and myeloid-differentiation-antigen-5 (MDA5) directly bind foreign dsRNA and assemble into filamentous signalling platforms using their RecA-like helicase domains [9] [10] [27] [28] [29]. However, ALRs are not RecA-like proteins, but consist of one oligomerization domain named pyrin domain (PYD) and one or two nonspecific dsDNA-binding HIN200 (hematopoietic interferon-inducible nuclear protein with 200 amino acids) domains (**Figure 3.1**) [30]. In contrast to Rig-I or MDA5 [9] [10] [27] [28] [29], the previous chapter's study of IFI16 suggests that nucleic acid binding and polymerization are allosterically coupled in ALRs.

There are two major unresolved questions regarding AIM2 inflammasome assembly. First, how AIM2 initiates the assembly is controversial. In contrast to the essential positive function of $IFI16^{PYD}$ in dsDNA binding and oligomerization [22], it was reported that the PYD of AIM2 ($AIM2^{PYD}$) plays an autoinhibitory function by blocking the dsDNA-binding surface of HIN200 domain ($AIM2^{Hin}$) [31] [32]. Although it was recently reported that isolated $AIM2^{PYD}$ auto-assembles into filaments [7] [33], whether the filamentation activity has any role in dsDNA binding remains unknown. Second, although it has been established that $AIM2^{PYD}$ directly induces the polymerization of ASC^{PYD} [7], how the recognition occurs at the

structural level remains speculative. For instance, many inflammatory signalling proteins contain PYDs and several highly conserved side chains that mediate multimeric PYD–PYD interactions have been identified [7] [33] [34]; however, only a subset of PYDs is known to interact with ASC [35] [36]. A recent study on foreign-dsRNA-sensing pathways provides a potential mechanism, as Hur and colleagues discovered that congruent multimeric architectures underpin the sequential oligomerization in the Rig-I signalling pathway [9]. For instance, Rig-I recruits the CARDs of mitochondrial antiviral-signalling protein ($MAVS^{CARD}$) into the helical 'oligomerization trajectory' of its CARD tetramers ($Rig\text{-}I^{CARD}$), consequently providing a structural template for the polymerization of the $MAVS^{CARD}$ filament [9]. Wu and colleagues have determined the architecture of the filament assembled by the PYD of ASC (ASC^{PYD}) [7]. However, because the architecture of $AIM2^{PYD}$ filament is unknown, it remains to be seen whether the AIM2–ASC axis operates in a similar manner.

Here it is found that AIM2 is not autoinhibited, but the size of dsDNA can act as a 'molecular ruler' to regulate the oligomerization of AIM2. Our collaborators also find that the helical symmetry of the $AIM2^{PYD}$ filament is consistent with that of ASC^{PYD}, suggesting that $AIM2^{PYD}$ filaments provide a template for assembling AIM2 filaments.

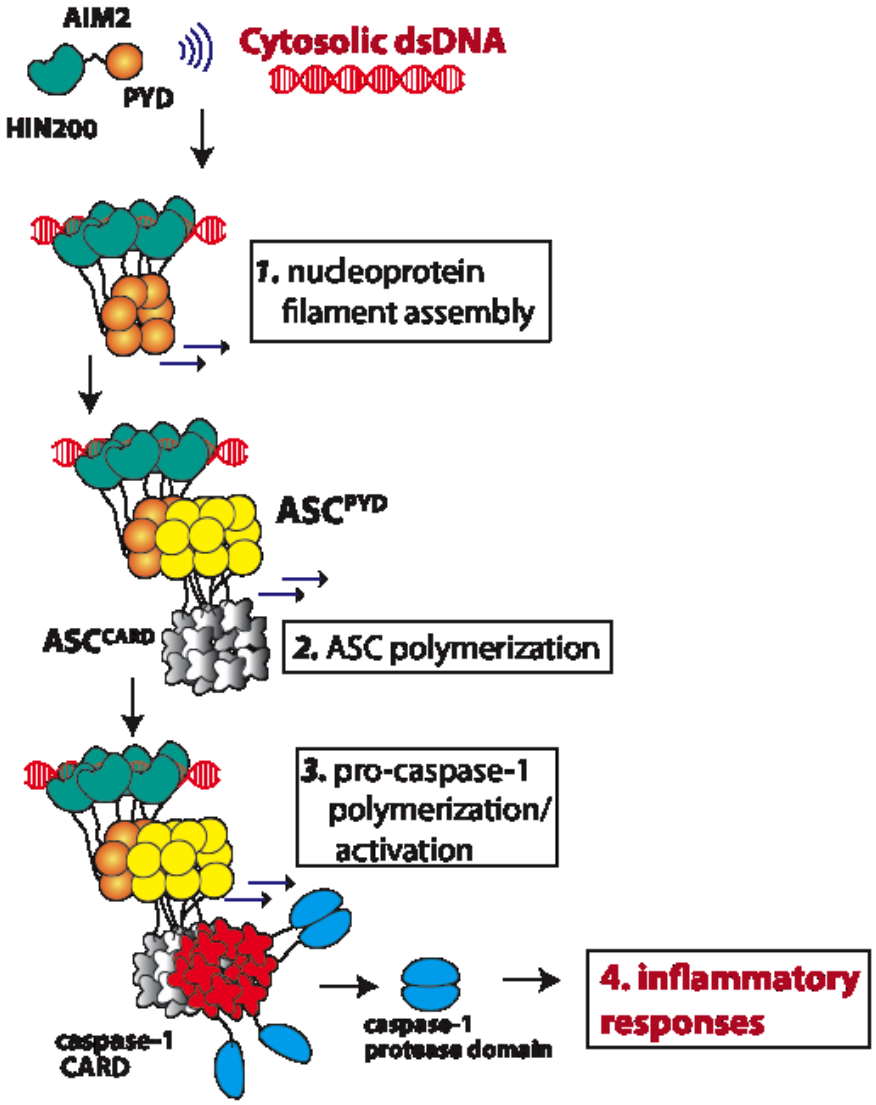

Figure 3.1: A model for the assembly and activation of the AIM2 inflammasome. AIM2 is comprised of an N-terminal PYD that oligomerizes and a C-terminal dsDNA-binding HIN200 domain, with an unstructured 50-amino acid linker between the two. ASC is a bipartite protein consisting of an N-terminal PYD and a C-terminal CARD. Procaspase-1 contains an N-terminal CARD followed by the protease domain. Only a few protease domains are shown for simplicity.

3.2 Results

3.2.1 N-Terminal MBP Masks the Oligomerization Activity of AIM2

No previous *in vitro* studies employed full-length AIM2 ($AIM2^{FL}$) without additional protein tags. Even the cornerstone of the autoinhibitory model proposed by Xiao and colleagues is an observation where N-terminal maltose-binding protein (MBP)-tagged full-length AIM2 ($AIM2^{FL}$) binds dsDNA more weakly than isolated $AIM2^{Hin}$ [31]. However, it was shown that N-terminal MBP masks the intrinsic oligomerization activity of PYDs by blocking a key interaction interface [7] [32] [33]. Thus, to re-examine the initiation of the AIM2 inflammasome assembly, recombinant $AIM2^{FL}$ and the isolated $AIM2^{Hin}$ in which N-terminal MBP tags are either left in place or removed by tobacco etch virus (TEV) protease were generated. The resulting AIM2 constructs all purify as monomers in gel filtration chromatography (**Figure 3.2A,B**). The binding affinity of MBP-tagged and -untagged AIM2 variants was compared by monitoring changes in fluorescence anisotropy of fluorescein amidite (FAM)-labelled 72-bp dsDNA (FAM-dsDNA72).

MBP-$AIM2^{FL}$ bound FAM-dsDNA72 nearly twofold more tightly than MBP-$AIM2^{Hin}$ (**Figure 3.3A** and **Table 3.1**). Without MBP, $AIM2^{FL}$ bound dsDNA72 at least 20-fold more tightly than isolated $AIM2^{Hin}$. Moreover, both tag-less AIM2 variants bound dsDNA72 significantly more tightly than their MBP-tagged counterparts (**Figure 3.3A**). Unlike the other AIM2 variants, the apparent binding constant of $AIM2^{FL}$ was near the concentration of ligand, suggesting that its binding to ligand is stronger than the lower limits of ligand concentrations used in the assays. These

data suggest that N-terminal MBP interferes with dsDNA binding of both $AIM2^{FL}$ and isolated $AIM2^{Hin}$, and that $AIM2^{PYD}$ plays a major positive role in dsDNA binding by AIM2.

Competition electrophoretic mobility shift assays with dsDNA fragments containing multiple $AIM2^{Hin}$-binding sites can test whether the initiation of the AIM2 inflammasome assembly on dsDNA proceeds via a two-state oligomerization binding mode [27] as opposed to the previously proposed non-interactive binding mechanism [31] [32]: no partially occupied intermediate dsDNA species are observed when a competitor is included in two-state oligomerization [27]. Here, both $AIM2^{FL}$ and isolated $AIM2^{Hin}$ showed an all-or-none transition; however, the MBP-tagged variants showed intermediates, further supporting a negative effect of MBP tags in the oligomerization-driven dsDNA binding of AIM2 (**Figure 3.3B,C**).

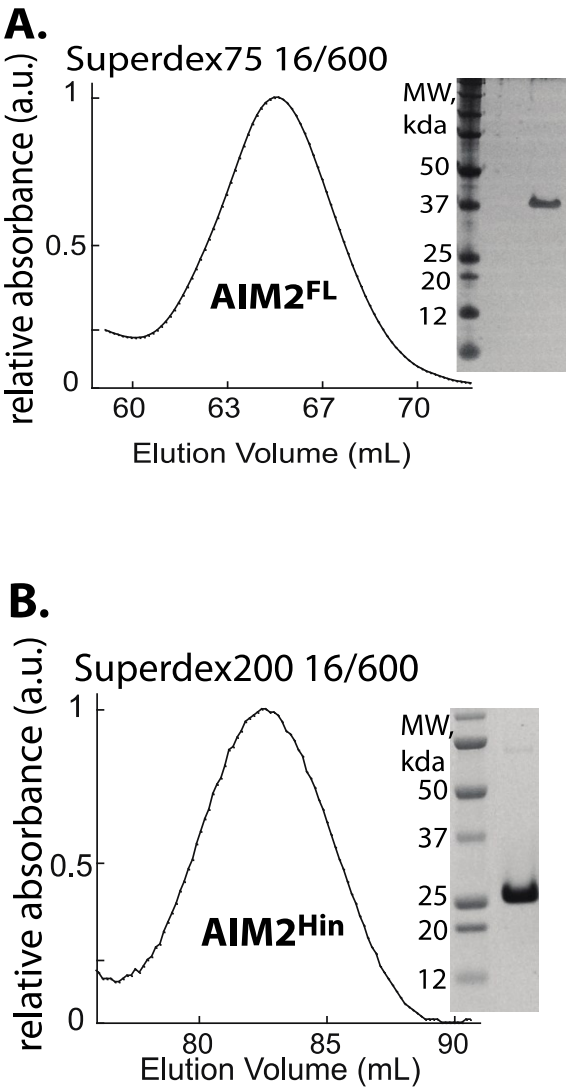

Figure 3.2: Purification of AIM2 constructs by gel-filtration chromatography. The column used is indicated above the sample trace. All elution volumes correspond to monomers according to the manufacturer's guidelines (GE Healthcare). (A) Sample UV-visible absorbance trace of AIM2 ^{FL}. To the right of the trace is a sample SDS-PAGE silver-stain gel of the combined peak fraction. (B) Sample UV-visible absorbance trace and silver-stained gel lane of AIM2 ^{Hin}.

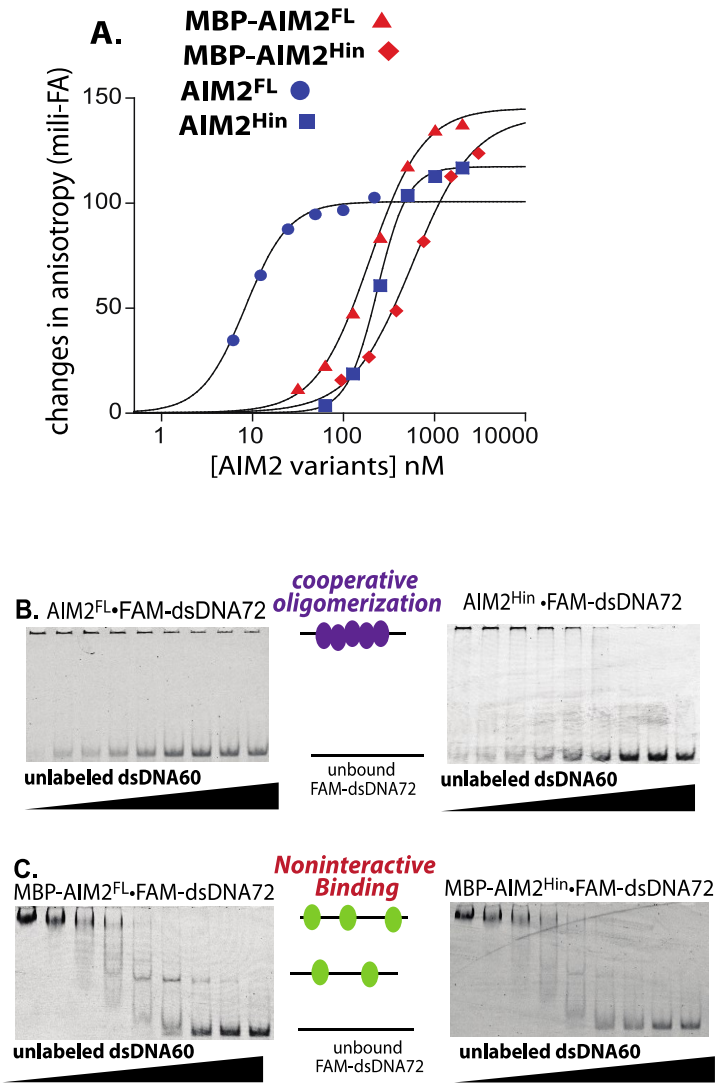

Figure 3.3: Oligomerization is integral to dsDNA binding by AIM2. (A) Binding of AIM2 variants to FAM-dsDNA72 (2.5nM) was monitored by changes in fluorescence anisotropy. The lines are fits to a Hill form of binding isotherm. The apparent binding constants (K_D) are determined by the Hill equation (bound = $1/(1 + K_{Dapp}/[IFI16])^{Hill\ Coefficient}$) and the values are listed in Table 3.1. All presented experiments were performed at least three times. (B,C) Competition EMSAs in which increasing concentrations of dsDNA60 (190, 95, 45, 23, 12, 6, 3, and 1.5μg/mL) were added to AIM2 (variants) FAM-dsDNA72 (200nM and 0.2μg/mL, respectively).

Variant	K_D (nM)	Hill Coefficient
MBP-AIM2^{FL}	234 ± 42	1.4 ± 0.1
MBP-AIM2 ^{Hin}	584 ± 22	1.1 ± 0.2
AIM2 ^{FL}	≤ 3	NA
AIM2 ^{Hin}	212 ± 28	1.7 ± 0.2
L10-11A-AIM2 ^{FL}	52 ± 11	1.7 ± 0.2
Acidic Patch to Ala	37 ± 12	1.4 ± 0.2
F27A-AIM2 ^{FL}	> 200	NA
T50A-AIM2 ^{FL}	28 ± 7	NA
K173A-AIM2 ^{FL}	113 ± 18	1.5 ± 0.3
Q258A-AIM2 ^{FL}	37 ± 6	ND
K272A-AIM2 ^{FL}	53 ± 13	1.7 ± 0.4
E147A-AIM2 ^{Hin}	≥ 1500	ND
F167A-AIM2 ^{Hin}	≥ 2000	ND
K173A-AIM2 ^{Hin}	≥ 2500	ND
K272A-AIM2 ^{Hin}	≥ 2500	ND

Table 3.1. AIM2 constructs binding to FAM-dsDNA72 (160 mM KCl). ± indicates standard deviation. n ≥ 3.

3.2.2 dsDNA Binding and Oligomerization are Integrated

When dsDNA-binding is dependent on oligomerization, the apparent affinity toward ligand changes with the size of dsDNA [13]; this provides a test for AIM2. For instance, the apparent binding affinity would increase nonlinearly up to the 'optimal' oligomer dictated by nucleic acid sizes [27]. To characterize the DNA-binding property of $AIM2^{FL}$ within the detection limit of the instruments, the reported salt concentration-dependent binding of $AIM2^{Hin}$ [31] was exploited and binding assays were performed with various dsDNA sizes at 400 mM KCl (**Figure 3.4**, **Table 3.2**). Even in this high salt condition, $AIM2^{FL}$ robustly bound FAM-dsDNA72 (**Figure 3.4A**). Importantly, $AIM2^{FL}$ bound larger dsDNA more tightly (**Figure 3.4A**), suggesting that oligomerization is integral to dsDNA binding. However, MBP-$AIM2^{FL}$, MBP-$AIM2^{Hin}$ and $AIM2^{Hin}$ all showed no detectable binding in this high salt condition (**Figure 3.4C**, **Table 3.3**), again supporting the positive role of $AIM2^{PYD}$ in dsDNA binding. In 160-mM KCl, isolated $AIM2^{Hin}$ also bound the larger dsDNA more tightly (**Figure 3.4B**, **Table 3.4**). Finally, both $AIM2^{FL}$ and isolated $AIM2^{Hin}$ bound the footprint-size dsDNA (10 bp) with minimal affinity (**Figure 3.4D**), suggesting that oligomerization is important for high affinity binding of both AIM2 variants (the dsDNA-binding footprint of $AIM2^{Hin}$ is ~9 bp [31] and that of $AIM2^{FL}$ is ~12 bp, **Figure 3.4E**).

Next, competition binding assays were performed to further investigate the relationship between dsDNA size and binding efficiency (dsDNA mass concentrations were used to normalize the number of available binding sites in each competitor; see **Table 3.5, Table 3.6, Table 3.7**). Consistent with the direct binding

data, both AIM2^{FL} and isolated AIM2^{Hin} bound larger dsDNA fragments significantly more tightly (**Figure 3.5A,B**). To allow isolated AIM2^{Hin} to bind dsDNA, 160mM KCl was used; the dsDNA size-dependent binding was apparent for AIM2^{FL} at both 400 and 160 mM KCl, confirming that buffer salt concentrations do not alter the overall mechanism (**Figure 3.5E,F**). The plots of half-maximal inhibition (IC_{50}) versus dsDNA size revealed a sigmoidal relationship where the difference in binding affinity between the near footprint-size dsDNA (dsDNA15) and dsDNA fragments larger than 300 bp can be as much as 4,000-fold for AIM2^{FL} and 150-fold for AIM2^{Hin}, respectively (**Figure 3.5C,D**). The cooperativity between the binding affinity and dsDNA size was assessed by using a Hill equation to fit the data (**Figure 3.5C,D**). The fitted Hill coefficient near four in **Figure 3.5C** suggests that the binding efficiency of AIM2^{FL} would improve 10,000-fold when the size of dsDNA is increased by only 10-fold. Furthermore, about 70-bp dsDNA was required to exit the lag phase ('threshold') and about 250- to 300-bp dsDNA fragments were required to achieve the 'optimal' efficiency for generating AIM2^{FL}·dsDNA complexes (**Figure 3.5C**). These dsDNA sizes in turn indicate that about 6 AIM2^{FL} molecules are required to assemble a 'threshold' oligomer and about 24 AIM2^{FL} molecules will generate an 'optimal' oligomer (**Figure 3.5C**). The observed dsDNA size-dependent binding also correlates with a previous *in vivo* observation in which the interleukin-1β secretion activity increased cooperatively between 10 and 80 bp transfected dsDNA [31]. Overall, the 'digitized' nucleoprotein complex-forming activity of AIM2 suggests that dsDNA can act as a 'molecular ruler' to control the assembly of the inflammasome in a switch-like mechanism.

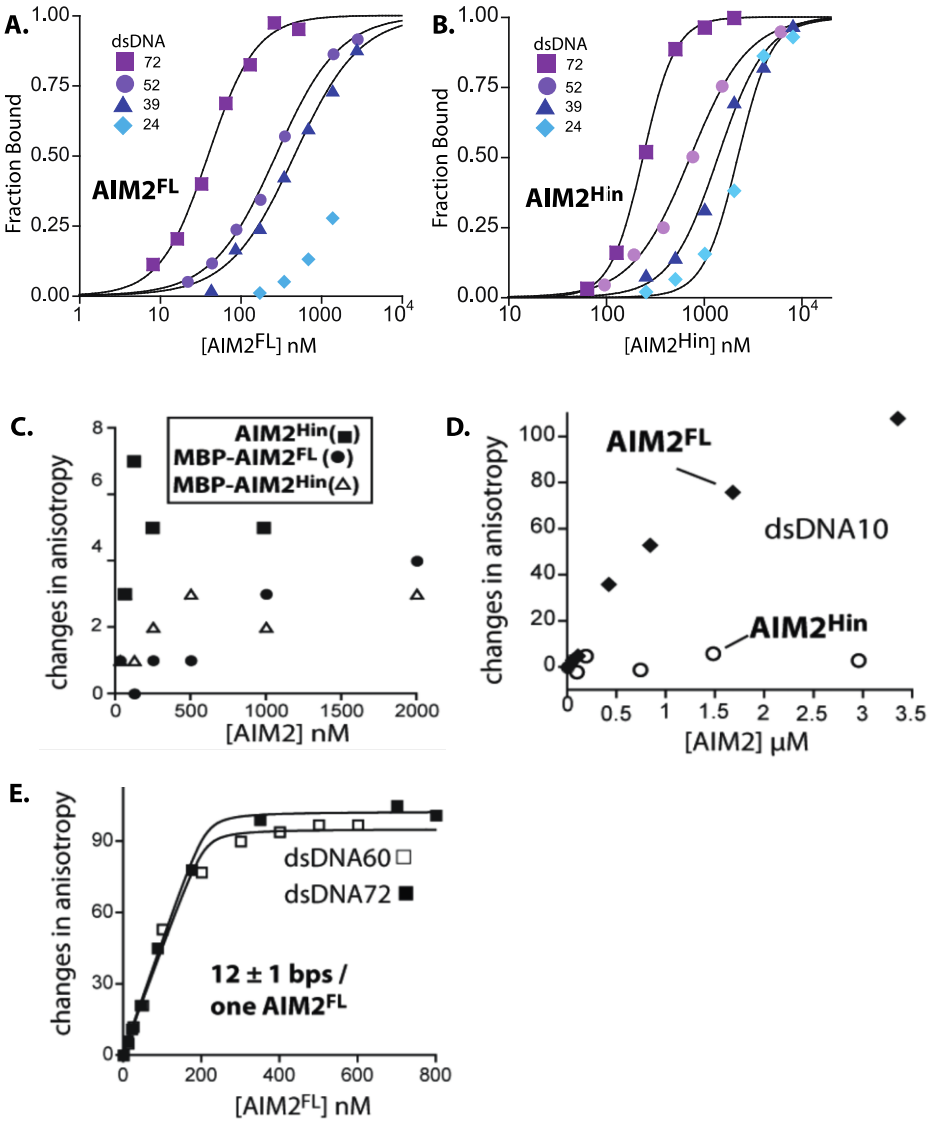

Figure 3.4: dsDNA binding characterization of AIM2^{FL} and isolated AIM2^{Hin}. (A) Binding of AIM2^{FL} and (B) isolated AIM2^{Hin} to each FAM-labeled dsDNA (1.5 nM) was determined by fluorescence anisotropy. The determined K_D values are listed in Table 3.2 and Table 3.4. (C) Titration of MBP-AIM2^{FL}, MBP-AIM2^{Hin}, and AIM2^{Hin} towards FAM-dsDNA72 (3 nM) in 400mM KCl. (D) Titration of AIM2^{FL} and isolated AIM2^{Hin} to FAM-dsDNA10 (5 nM) in 160mM KCl. (E) Stoichiometric titration of AIM2^{FL} to 40 nM FAM-dsDNA60 and dsDNA72 in 160mM KCl to determine its footprint. The inflection points were 210 and 213 nM for dsDNA60 and dsDNA72, respectively.

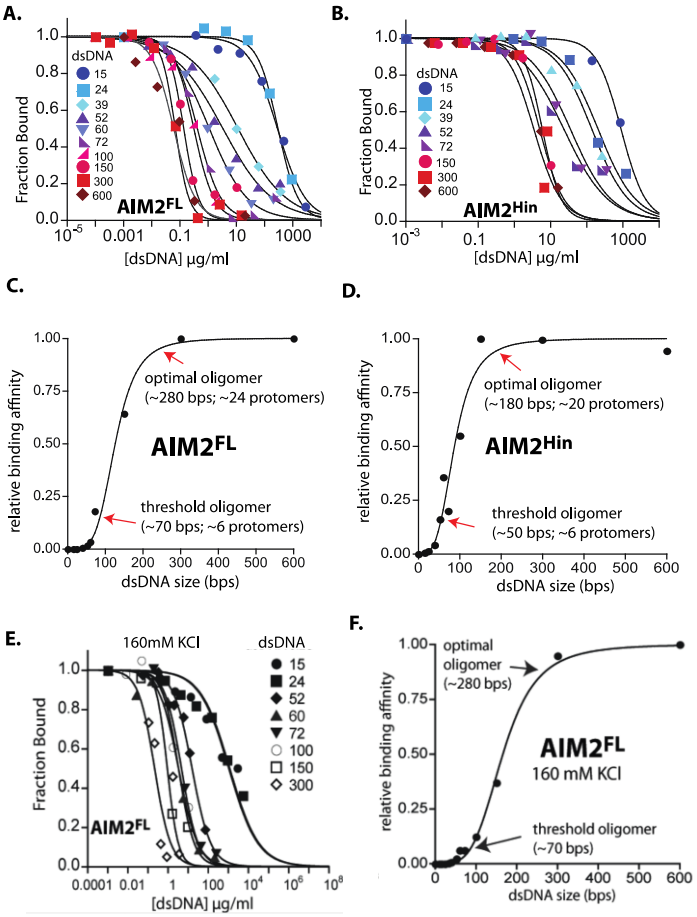

Figure 3.5: AIM2^{FL} and isolated AIM2^{Hin} bind dsDNA in a length-dependent manner. (A) Competition binding assays using FAM-dsVACV72 (1.5 nM, 0.06 µg/ml) and AIM2^{FL} (70 nM) at 400 mM KCl against various dsDNA fragments; the lines are fits to a competition binding equation: 1/(1+([DNA$_{competitor}$]/IC$_{50}$)^Hill Coefficient). The determined IC$_{50}$ values are listed in Table 3.5. (B) Competition binding assays using FAM-dsVACV72 (5 nM, 0.2µg/ml) and AIM2^{Hin} (250 nM) at 160 mM KCl against various DNA fragments. The determined values are listed in Table 3.6. The plots of the binding efficiency vs. the length of dsDNA for AIM2^{FL} (C) and AIM2^{Hin} (D). The binding efficiency was determined by normalizing the mean IC$_{50}$ of each fragment to that of dsDNA600, and the data were fit to the Hill equation (the Hill coefficient for (C) is 4.2 ± 0.2 and (D) is 3.7 ± 0.3). The "threshold" oligomer is defined as the size of dsDNA (AIM2 cluster) required to exit the apparent lag phase, and the "optimal" oligomer is the size of dsDNA (AIM2 cluster) required to reach the inflection point. (E) Competition binding assays of AIM2^{FL} (250 nM) and FAM-dsDNA72 (5 nM) against various dsDNA at 160 mM KCl. The determined K$_D$ are listed in Table 3.7. (F) The relative binding efficient of AIM2^{FL} with respect to the size of dsDNA from (E) was determined as in (C).

DNA Length (bp)	K_D (nM)	Hill Coefficient
24	≥ 2000	ND
39	531 ± 93	1.2 ± 0.1
52	350 ± 12	1.1 ± 0.2
72	39 ± 12	1.5 ± 0.3

Table 3.2: Binding of $AIM2^{FL}$ towards various dsDNA lengths (400 mM KCl). ± indicates standard deviation. n ≥ 3.

Variant	K_D (nM)	Hill Coefficient
MBP-$AIM2^{FL}$	NB	ND
MBP-$AIM2^{Hin}$	NB	ND
$AIM2^{FL}$	39 ± 12	1.5 ± 0.3
$AIM2^{Hin}$	NB	ND
L10-11A-$AIM2^{FL}$	NB	ND
Acidic Patch to Ala	NB	ND
F27A-$AIM2^{FL}$	NB	ND
I46D- $AIM2^{FL}$	NB	ND
T50A-$AIM2^{FL}$	95 ± 31	1.4 ± 0.1
E147A-$AIM2^{FL}$	143 ± 45	1.4 ± 0.2
F167A-$AIM2^{FL}$	ND	ND
K173A-$AIM2^{FL}$	ND	ND
Q258A-$AIM2^{FL}$	31 ± 5	1.4 ± 0.2

Table 3.3. AIM2 constructs binding to FAM-dsDNA72 (400 mM KCl). ± indicates standard deviation. n ≥ 3.

DNA Length (bp)	K_D (nM)	Hill Coefficient
10	>> 2000	ND
24	2110 ± 486	2.1 ± 0.4
39	1205 ± 339	2.2 ± 0.5
52	922 ± 168	1.4 ± 0.2
72	212 ± 33	2.1 ± 0.4

Table 3.4: Binding of $AIM2^{Hin}$ towards various dsDNA lengths (160 mM KCl). ± indicates standard deviation. n ≥ 3.

DNA Length (bp)	IC_{50} (μg/mL)	Hill Coefficient
15	1695 ± 367	0.74 ± 0.17
24	1608 ± 151	0.69 ± 0.26
52	16 ± 6	0.97 ± 0.04
60	6 ± 2	1.00 ± 0.10
72	6.1 ± 1.8	1.17 ± 0.12
100	3.1 ± 0.3	0.8 ± 0.3
150	1.2 ± 0.3	1.5 ± 0.4
300	0.39 ± 0.24	2.24 ± 0.16

Table 3.5: $AIM2^{FL}$ competition binding (160 mM KCl). ± indicates standard deviation. n ≥ 3.

DNA Length (bp)	IC_{50} (μg/mL)	Hill Coefficient
15	805 ± 118	1.3 ± 0.1
24	365 ± 87	0.8 ± 0.3
39	119 ± 32	0.7 ± 0.2
52	31 ± 5	0.6 ± 0.2
60	14 ± 1	0.9 ± 0.4
72	25 ± 1	0.7 ± 0.1
100	9.1 ± 0.7	1.2 ± 0.7
150	4.1 ± 1.8	1.3 ± 0.3
300	4.8 ± 0.5	1.5 ± 0.3
600	5.3 ± 0.2	1.3 ± 0.1

Table 3.6: $AIM2^{Hin}$ competition binding (160 mM KCl). ± indicates standard deviation. n ≥ 3.

DNA Length (bp)	IC_{50} (μg/mL)	Hill Coefficient
15	328 ± 39	0.9 ± 0.4
24	237 ± 67	0.8 ± 0.4
39	16 ± 4	0.7 ± 0.3
52	5.2 ± 0.5	0.8 ± 0.2
60	2.6 ± 0.3	0.8 ± 0.1
72	0.5 ± 0.1	0.8 ± 0.4
100	0.29 ± 0.04	1.2 ± 0.1
150	0.14 ± 0.01	1.5 ± 0.1
300	0.09 ± 0.02	1.6 ± 0.1
600	0.09 ± 0.02	1.4 ± 0.3

Table 3.7: $AIM2^{FL}$ Competition Binding (400 mM KCl). ± indicates standard deviation. n ≥ 3.

3.2.3 Mutagenesis Studies Support the Positive Role of AIM2^{PYD}

In the previous chapter, it was found that several highly conserved surface side chains of IFI16 mediate its oligomerization-driven dsDNA-binding mechanism (**Figure 3.6A**). For AIM2, several equivalently positioned surface side chains mediate the auto-oligomerization of isolated AIM2^{PYD} *in vivo* [33] and interaction with ASC^{PYD} *in vitro* [7](for example, Leu10, Leu11 and Phe27 in **Figure 3.6B**). However, whether these side chains have any role in dsDNA binding remains unknown. To further test the positive role of AIM2^{PYD} in dsDNA binding, a panel of surface residue variants based on the study of IFI16 as well as sequence conservation (**Figure 3.6A,B**; none of these side chains are implicated in the alleged autoinhibition of AIM2 [31] [32], see also below) was generated. Almost all of these PYD variants significantly disrupted the dsDNA-binding activity of AIM2^{FL} (**Figure 3.6C**), supporting that oligomerization of PYD plays a major positive function in dsDNA binding.

Xiao and colleagues proposed that a unique acidic surface of AIM2^{PYD} not present in other related ALR docks to the basic dsDNA-binding surface of AIM2^{Hin}, thus stabilizing the autoinhibited conformation in the absence of dsDNA [31] [32] (**Figure 3.6A,B**; designated 'acidic patch'; Asp19, Glu20, Glu21 and Asp23). Such an autoinhibitory model entails that neutralizing the acidic patch would allow AIM2^{FL} to bind dsDNA more tightly. However, D19A-E20A-E21A-D23A-AIM2^{FL} failed to bind FAM-dsDNA72 (**Figure 3.6C**). This observation disagrees with the inhibitory role of the acidic patch, but is again consistent with the idea that AIM2^{PYD} plays a positive role in dsDNA binding.

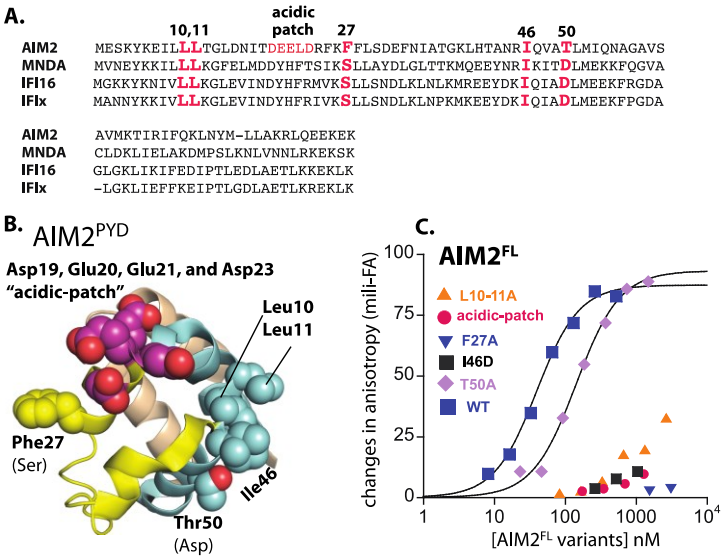

Figure 3.6: Mutagenesis studies to test the role of AIM2^{PYD} in dsDNA binding. (A) The sequence alignment of ALRs. The side-chains mutated here are indicated in red. (B) The crystal structure of AIM2^{PYD} (PDB ID: 3VD8). The mutated side-chains are shown as spheres. The amino acids indicated in the parentheses are the equivalent IFI16 residues. (C) Binding of FAM-dsDNA72 (1.5 nM) by various AIM2^{FL} PYD-mutants were tested at 400 mM KCl. The determined K_D values are listed in Table 3.3.

3.2.4 The oligomerization of AIM2^{Hin} is shared with murine p202

In contrast to the dsDNA-binding HIN200 domains of IFI16, these dsDNA-binding studies suggest that AIM2^{Hin} alone can oligomerize on dsDNA (**Figure 3.4B** and **Figure 3.5B,D**). On the other hand, a murine ALR named p202 inhibits the activity of AIM2 by binding to AIM2^{Hin} with its tetrameric second HIN200 domain (p202^{HinB}); p202^{HinB} does not bind dsDNA [37]. Several p202^{HinB} side chains implicated in tetramerization [37] are conserved in AIM2^{Hin} (**Figure 3.7A,B**). To test whether AIM2^{Hin} uses similarly positioned side chains as p202^{HinB} to cluster on dsDNA, the indicated side chains in **Figure 3.7A,B** were mutated on both AIM2^{FL} and isolated AIM2^{Hin}. Mutations distal to the dsDNA-binding surface significantly decreased dsDNA binding by both AIM2^{FL} and AIM2^{Hin} (**Figure 3.7C,D**). These results not only support the idea that the oligomerization of AIM2^{Hin} is important for dsDNA binding but also suggest that the oligomerization of AIM2^{Hin} is an evolutionarily conserved feature.

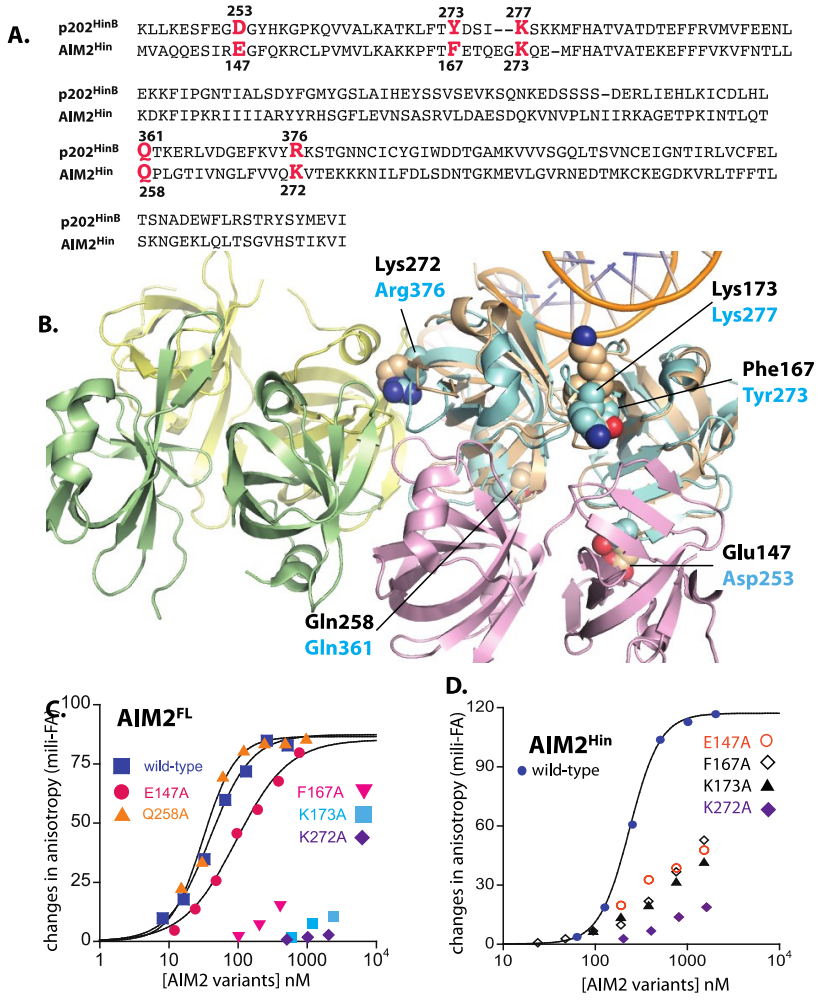

Figure 3.7: Mutagenesis studies to test the evolutionarily conserved oligomerization activity AIM2^{Hin} in dsDNA binding. (A) The sequence alignment of AIM2^{Hin} and p202^{HinB}. The side-chains mutated here are indicated in red. (B) The crystal structure of dsDNA-bound AIM2^{Hin} (colored in "wheat;" PDB ID: 3RN2) aligned to the tetramer structure of p202^{HinB} (PDB ID: 4L5T); root mean squared deviation of alignment is 1.5Å. The p202^{HinB} protomers are colored in green, yellow, pink, and cyan, respectively. The mutated side-chains are shown as spheres. AIM2 side-chains are labeled in black and those from p202 are indicated in cyan. (C) Binding of FAM-dsDNA72 (1.5 nM) by various AIM2^{FL} HIN200-mutants were tested at 400 mM KCl. The determined K_D values are listed in Table 3.3. (D) Binding of various AIM2^{Hin} mutants were tested on FAM-dsDNA72 (5 nM) at 160 mM KCl. The determined K_D values are listed in Table 3.1.

3.2.5 AIM2^{PYD} is required to oligomerize when dsDNA is in excess

When basal AIM2 encounters foreign dsDNA in the cytoplasm, the individual molecules must assemble into the inflammasome even in the presence of excess binding sites (for example, nearly 400,000 binding sites are in the genome of one *F. tularensis*). To test this idea, two separate AIM2^{FL} populations were labeled with a Förster resonance energy transfer (FRET) donor and acceptor (**Figure 3.8A**, top). As previously observed for full-length IFI16, saturating FRET signals were indeed detected from AIM2^{FL} in a dsDNA size-dependent manner even when the substrate is present in excess (**Figure 3.8A,B**; **Table 3.8**). FRET signals were detected from dsDNA fragments as short as 24 bp, suggesting that the minimal binding unit for AIM2 oligomerization is a dimer compared with a tetramer for IFI16 (see previous chapter). Consistent with the competition assays (**Figure 3.5**), plotting the normalized binding efficiency versus dsDNA length from the FRET assay data also suggests a cooperative relationship in which the binding affinity of AIM2^{FL} can increase as much as 1,000-fold when the size of available dsDNA is 10 times longer (the Hill coefficient is ~3; **Figure 3.8C**). FRET signals from labelled AIM2^{Hin} were only detected if the salt concentration of the reaction buffer was lowered to 60 mM KCl (**Figure 3.8D**). Unlike AIM2^{FL}, FRET signals from labelled AIM2^{Hin} peaked at the dsDNA concentration equivalent to the amount of AIM2^{Hin} present in these assays, but decreased with excess dsDNA (**Figure 3.8D**). Also, unlike AIM2^{FL}, the peak amplitude was correlated with the size of each dsDNA (**Figure 3.8D**). These observations suggest that the AIM2^{Hin} oligomers are likely different from those assembled by AIM2^{FL}, and that AIM2^{PYD} is required for robust dsDNA binding and

polymerization in the presence of excess dsDNA. In addition, the observed cooperative relationship between dsDNA size and oligomerization activity (**Figure 3.8C**) is consistent not only with the competition experiments (**Figure 3.5C**) but also with the previous *in vivo* observation [31], thus further supporting a 'digital ruler' concept in the regulation of the AIM2 inflammasome.

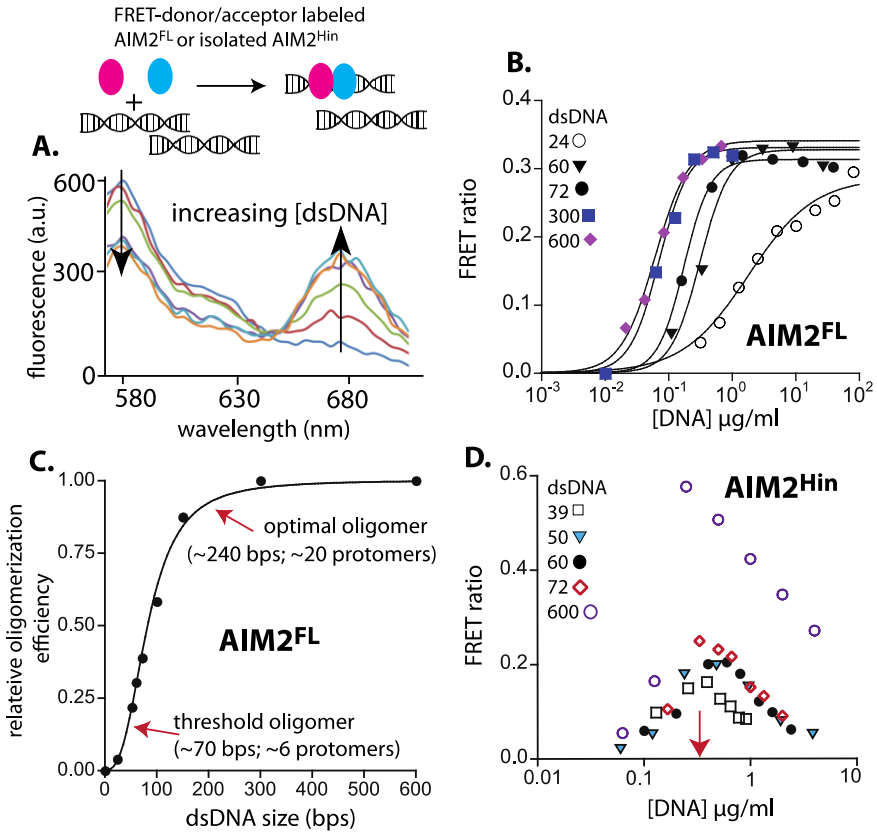

Figure 3.8: AIM2^{PYD} is necessary for oligomerization and dsDNA binding in the presence of excess dsDNA. (A) Top: a cartoon demonstrating the rationale of the described FRET experiments. The two differentially colored ovals represent fluorophore (Dylight-550 and Dylight-650) labeled AIM2. Bottom: A sample fluorescence emission spectra of an equimolar mixture of FRET donor and acceptor labeled AIM2FL. A.U., arbitrary units. (B) Changes in the ratio between the FRET donor emission (λ_{max}: 578 nm) and the acceptor emission (λ_{max}: 678 nm) at each indicated dsDNA concentration. The apparent oligomerization constants (K_{DF}) were obtained by fitting the data to a Hill equation and are listed in Table 3.8. (C) A plot of binding efficiency vs. the length of dsDNA for AIM2^{FL}. The data were fit to the Hill equation (the Hill coefficient is 3.2 ± 0.3). The efficiency was determined by normalizing the mean K_{DF} of each fragment to that of dsDNA600. (D) The FRET ratio of AIM2^{Hin} with increasing amounts of various dsDNA. The red arrow indicates the concentration of AIM2^{Hin} present in the assay.

DNA Length (bp)	K_D (µg/mL)	Hill Coefficient
24	1.75 ± 0.57	0.9 ± 0.1
39	0.16 ± 0.01	1.4 ± 0.4
52	0.32 ± 0.01	1.9 ± 0.3
60	0.23 ± 0.06	1.6 ± 0.3
72	0.18 ± 0.02	2.1 ± 0.4
100	0.12 ± 0.02	1.6 ± 0.1
150	0.08 ± 0.01	1.5 ± 0.2
200	0.07 ± 0.02	1.5 ± 0.2
300	0.07 ± 0.02	1.2 ± 0.2
600	0.07 ± 0.01	2.6 ± 0.7

Table 3.8: AIM2^{FL} FRET data (160mM KCl). ± indicates standard deviation. n ≥ 3.

3.2.6 apo-AIM2^{FL} can auto-oligomerize

The autoinhibitory model entails that dsDNA is required to unlock monomeric AIM2 and initiate oligomerization [31] [32]. By contrast, because untagged AIM2^{FL} binds dsDNA much more tightly than AIM2^{Hin}, the role of dsDNA may be to increase the local concentration of AIM2 by acting as a 'one-dimensional ruler,' consequently improving the prospects for forming AIM2^{PYD}·AIM2^{PYD} encounter complexes. This model entails that dsDNA-free AIM2^{FL} should be able to oligomerize on its own if the concentration threshold is met. On the other hand, *apo*-AIM2^{FL} should remain monomeric according to the autoinhibitory model. To test these opposing predictions, negative stain electron microscopy (ns-EM) was used to probe the oligomeric state of *apo*-AIM2^{FL} at various concentrations. Disagreeing with the autoinhibitory model, *apo*-AIM2^{FL} forms filaments (**Figure 3.9A**) in a protein concentration-dependent manner (≥1 μM; **Figure 3.9D**). Closer inspection of the electron micrographs revealed that DNA-free AIM2^{FL} filaments assume 'Brussels Sprout'-like structures in which the central filaments ('core stems') are decorated with speck-like densities at the periphery ('sprouts') (**Figure 3.9C**). Although the core stem appeared well ordered, the peripheral specks appeared random and often disordered (**Figure 3.9A,C,D,E**). On the basis of the reported auto-filamentation activity of isolated AIM2^{PYD} [7] [33], it is likely that AIM2^{PYD} forms the core stem and the AIM2^{Hin} clusters are flexibly attached via the intrinsically disordered linker region (50 amino acids); the 9-nm diameter of the core stem also corresponds to the width of the ASC^{PYD} filament [7]. In addition, no auto-assembled AIM2^{Hin} filaments were observed (**Figure 3.9B**). AIM2^{FL} filaments

were detected even at 1.6 M NaCl, thus suggesting that they are more resilient against the environment than the ASC^{PYD} filament [7] (**Figure 3.9E**). Considering this auto-oligomerization activity, the concentrations of wild-type $AIM2^{FL}$ were kept as low as possible in the biochemical assays (typically <100 nM; **Figures 3.3-3.8**). Moreover, saturating binding isotherms were observed in most of the assays (**Figure 3.3A**), suggesting that auto-oligomerization did not cause any significant artifacts.

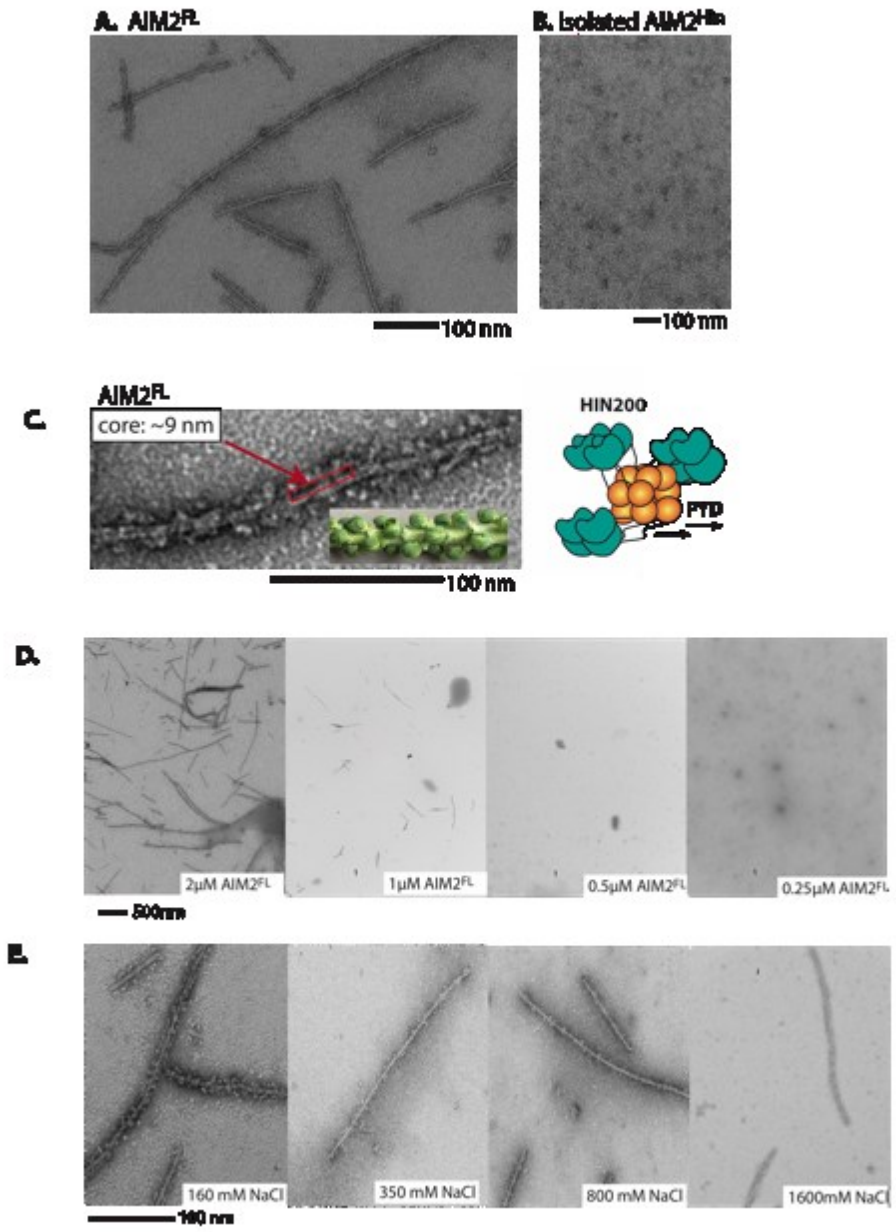

Figure 3.9: AIM2^{FL} assembles into filaments without dsDNA. (A) A negatively stained electron micrograph of AIM2^{FL} at 2 μM. (B) A negatively stained electron micrograph of AIM2^{Hin} at 5 μM. (C) Higher magnification of the AIM2 filament. The inset is unpicked Brussels sprout, and the cartoon on the right is the proposed overall arrangement of the filament. The red box in C indicates the stable "core stem" of the AIM2^{FL} filament. (D) Electron micrographs of AIM2^{FL} at various concentrations. The scale bar is 500 nm. (E) Electron micrographs of AIM2^{FL} (2 μM) at various salt concentrations. The scale bar is 100 nm.

3.2.7 dsDNA-binding deficient variants fail to auto-oligomerize

If the AIM2 variants identified in the present study indeed affect oligomerization, but not direct protein–dsDNA interactions, they should also disrupt the auto-filamentation activity. Hence, these variants were subjected to examination using ns-EM. The Brussels sprout-like filaments were detected from the $AIM2^{FL}$ variants that lacked any major defects in dsDNA binding (**Figure 3-10**). However, the $AIM2^{FL}$ variants with impaired dsDNA binding did not show any auto-assembled filaments regardless of whether a mutation is located on the PYD or HIN200 domains (**Figure 3-10**). The failure to auto-assemble by mutating the PYD side chains suggests that it is the filamentation activity of $AIM2^{PYD}$ that positively contributes to dsDNA binding. Moreover, the lack of auto-oligomerization resulting from mutating the HIN200 domain side chains suggests that although the oligomerization activity of $AIM2^{PYD}$ outweighs that of $AIM2^{Hin}$, both domains have positive functions in assembling $AIM2^{FL}$ filaments in the presence or absence of dsDNA.

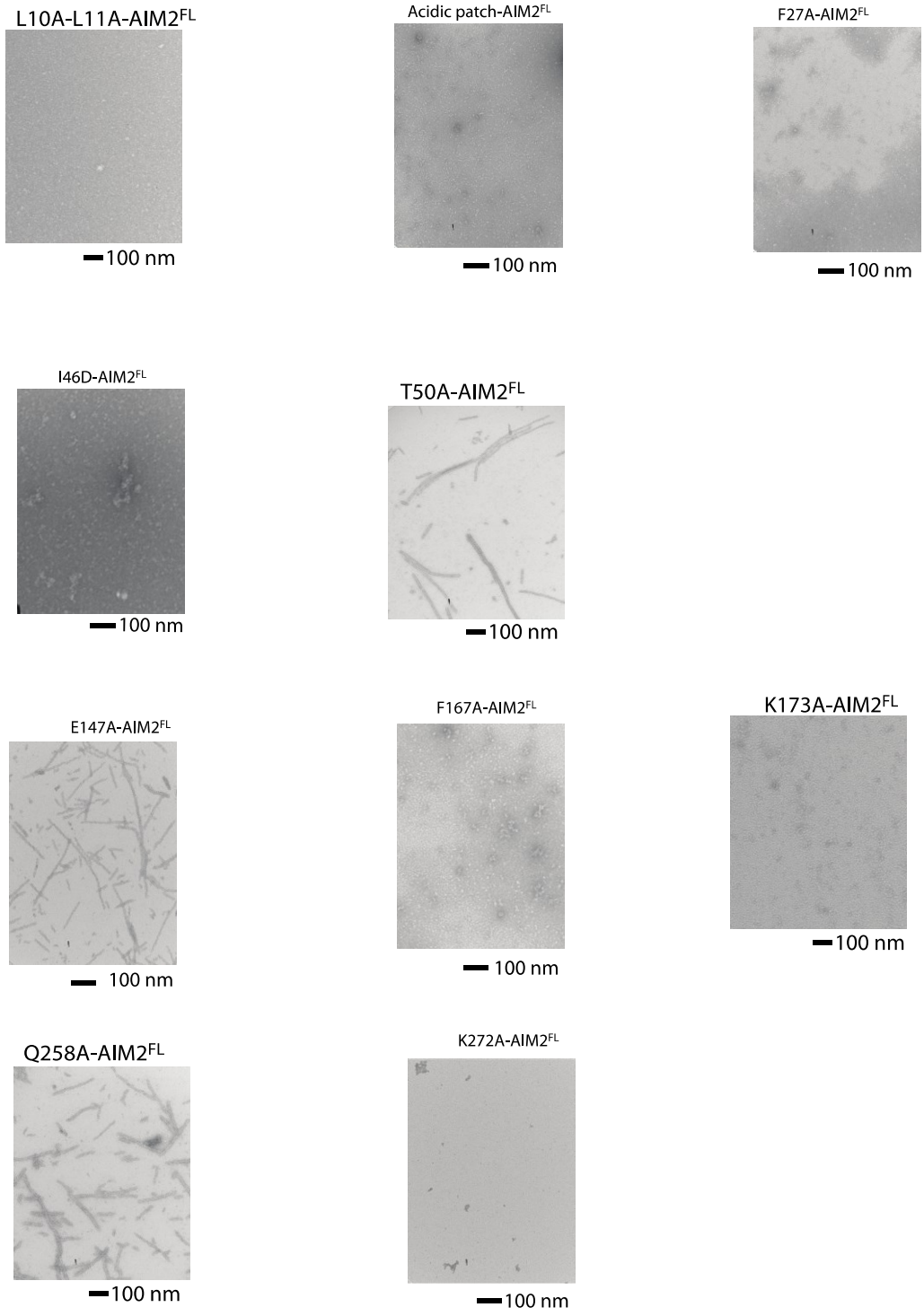

Figure 3.10: Select mutations disrupt AIM2 auto-filaments. Shown are electron micrographs of $AIM2^{FL}$ variants at 2 μM.

3.2.8 Helical symmetry of the AIM2^{PYD} filament

Several previous studies reported the direct interaction between AIM2^{PYD} and ASC^{PYD} [7] [17] [20]. The helical architecture of the ASC^{PYD} filament is known [7]; however, whether or not the upstream AIM2 oligomer provides a 'polymerization template ' via a congruent oligomeric architecture as seen from the Rig-I·MAVS interaction [9] is an open question. Thus, our collaborator Edward Egelman determined the helical symmetry of the core stem of the AIM2^{FL} filaments. The average power spectrum of the AIM2^{FL} filaments obtained from ns-EM is remarkably similar to that of the ASC^{PYD} filament (**Figure 3.11A**, see also ref. [7]), which is an ~90-Å wide, six-start helix with three-fold symmetry [7]. These parameters also agree with the proposition in which the AIM2^{PYD} filament constitutes the core stem (**Figure 3.9C**). On the basis of the consistent helical symmetry, a homology model of the AIM2^{PYD} filament was generated using the cryo-EM structure of the ASC^{PYD} filament as a template [7], where it was found that all the side chains identified to be important for dsDNA binding and auto-oligomerization are located at the subunit interfaces (**Figure 3.11B,C**). Overall, the EM analysis suggests that the upstream AIM2^{PYD} filament provides a structural template for the polymerization of downstream ASC^{PYD} (**Figure 3.1**).

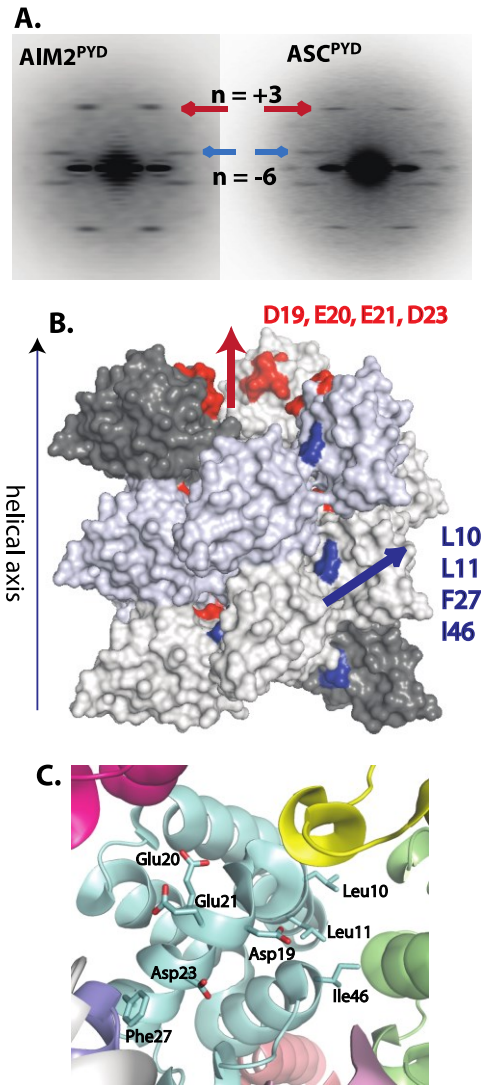

Figure 3-11: The congruent helical symmetry between filaments assembled by AIM2^{PYD} and ASC^{PYD}. (A) The ns-EM average power-spectra of the AIM2^{PYD} filament (left) and the ASC^{PYD} filament (right; unpublished, courtesy of Dr. Hao Wu, Harvard). The colored arrows indicate corresponding helical symmetry lines observed from both filaments. (B and C) A homology of model of the AIM2^{PYD} filament based on the cryo-EM structure of the ASC^{PYD} filament (PDB ID: 3J63). The AIM2 side-chains important for dsDNA binding and auto-oligomerization are highlighted.

3.2.9 PYD interactions dictate the filament architecture

A previous cell-based imaging study showed that isolated $AIM2^{PYD}$ and $AIM2^{FL}$ form filamentous aggregates, but $AIM2^{Hin}$ failed to form such structures [33]; whether dsDNA (transfected plasmid) is part of the filamentous $AIM2^{FL}$ aggregates is unknown. The present study is consistent with this *in vivo* observation [33], as dsDNA-free $AIM2^{FL}$ assembles into filaments via its PYD. However, in principle, the HIN200 domains of $AIM2^{FL}$ could bind dsDNA along its length, and thus might also generate an ordered filamentous structure. Thus, to further resolve whether $AIM2^{PYD}$ or $AIM2^{Hin}$ oligomers dictate the overall architecture of dsDNA-bound $AIM2^{FL}$, the morphologies of $AIM2^{FL}$ and $AIM2^{Hin}$ bound to λ-phage dsDNA (λdsDNA) were determined using ns-EM. Isolated $AIM2^{Hin}$ did not show any ordered filaments, but displayed random 'beads on a string'-like clusters on λdsDNA (**Figure 3.12A**). By contrast, the Brussels sprout-like filaments were no longer detected on adding dsDNA to $AIM2^{FL}$, but new larger filaments about two- to three-times wider than the DNA-free filaments appeared in the micrographs (20–25 nm; **Figure 3.12B**), indicating dsDNA binding. Together, these observations suggest that $AIM2^{Hin}$ binds dsDNA and clusters randomly, and that $AIM2^{PYD}$ oligomers underpin the filamentous architecture of dsDNA-bound $AIM2^{FL}$.

λdsDNA is about 50 kb and displayed random-coil structures (**Figure 3.12B**); however, all observed dsDNA-bound $AIM2^{FL}$ filaments seemed well ordered, further suggesting that PYD–PYD interactions, but not $AIM2^{Hin}$–dsDNA interactions (**Figure 3.12A,B**), dictate the overall architecture of λdsDNA-bound $AIM2^{FL}$ filaments. Because unbound λdsDNA can be observed in the $AIM2^{FL}$·dsDNA samples (**Figure**

3.12B), the EM experiments also reveal that AIM2^{FL} filaments assemble from random positions on dsDNA and that the filaments can assemble even in the presence of excess dsDNA (**Figure 3.12B**). These results not only corroborate the FRET assays in which AIM2^{FL} oligomerized in the presence of excess dsDNA (**Figure 3.8**) but also are consistent with the oligomerization activity of IFI16, which also formed filaments in the presence of excess dsDNA via its PYD (see previous chapter). Several filaments apparently merged laterally and became intertwined, likely reflecting the punctate-like AIM2 inflammasome structures observed from *in vivo* studies [20] [31] (**Figure 3.12C,D**).

To further test that AIM2^{PYD} oligomers, but not those assembled by AIM2^{Hin}, underpin the filamentous structure of dsDNA-bound AIM2^{FL}, the morphologies of several λdsDNA-bound AIM2^{FL} variants with defective dsDNA-binding/auto-assembly activities were examined. Conditions that allow these AIM2^{FL} variants to bind dsDNA were identified (160 mM KCl; same as all wild-type EM experiments); all the defective variants still bound dsDNA significantly more weakly than wild type (**Table 3.1**). Under these conditions, disrupting the oligomerization activity of AIM2^{Hin} still allowed the full-length protein to assemble into isomorphic filaments on λdsDNA as wild type (**Figure 3.12E–G**). The I46D mutation located in the PYD also resulted in wild type-like AIM2^{FL} filaments once bound to λdsDNA (**Figure 3.12H**). However, L10A-L11A-AIM2^{FL} showed a heterogeneous nucleoprotein filament population (**Figure 3.12I**). More strikingly, D19A-E20A-E21A-D23A-AIM2^{FL} (acidic patch) displayed disordered clusters similar to λdsDNA-bound isolated AIM2^{Hin} (**Figure 3.12J**). The homology model indicates that Leu10, Leu11

and Ile46 mediate the radial interactions in the AIM2^{PYD} filament, while the acidic patch mediates the axial interactions (**Figure 3.11B** and **Figure 3.12K**). Thus, these results suggest that the dsDNA scaffold can at least partially restore the filamentous structure if the radial interactions in the AIM2^{PYD} filament are compromised (**Figure 3.11B** and Figure **3.12K**). However, disrupting the axial interactions seems to be detrimental to the filamentous structure even when bound to the dsDNA scaffold (**Figure 3.11B** and **Figure 3.12K**). In reviewing the previous experiments, it appears that filamentous AIM2^{PYD} oligomers underpin the architecture of dsDNA-bound AIM2^{FL} polymers.

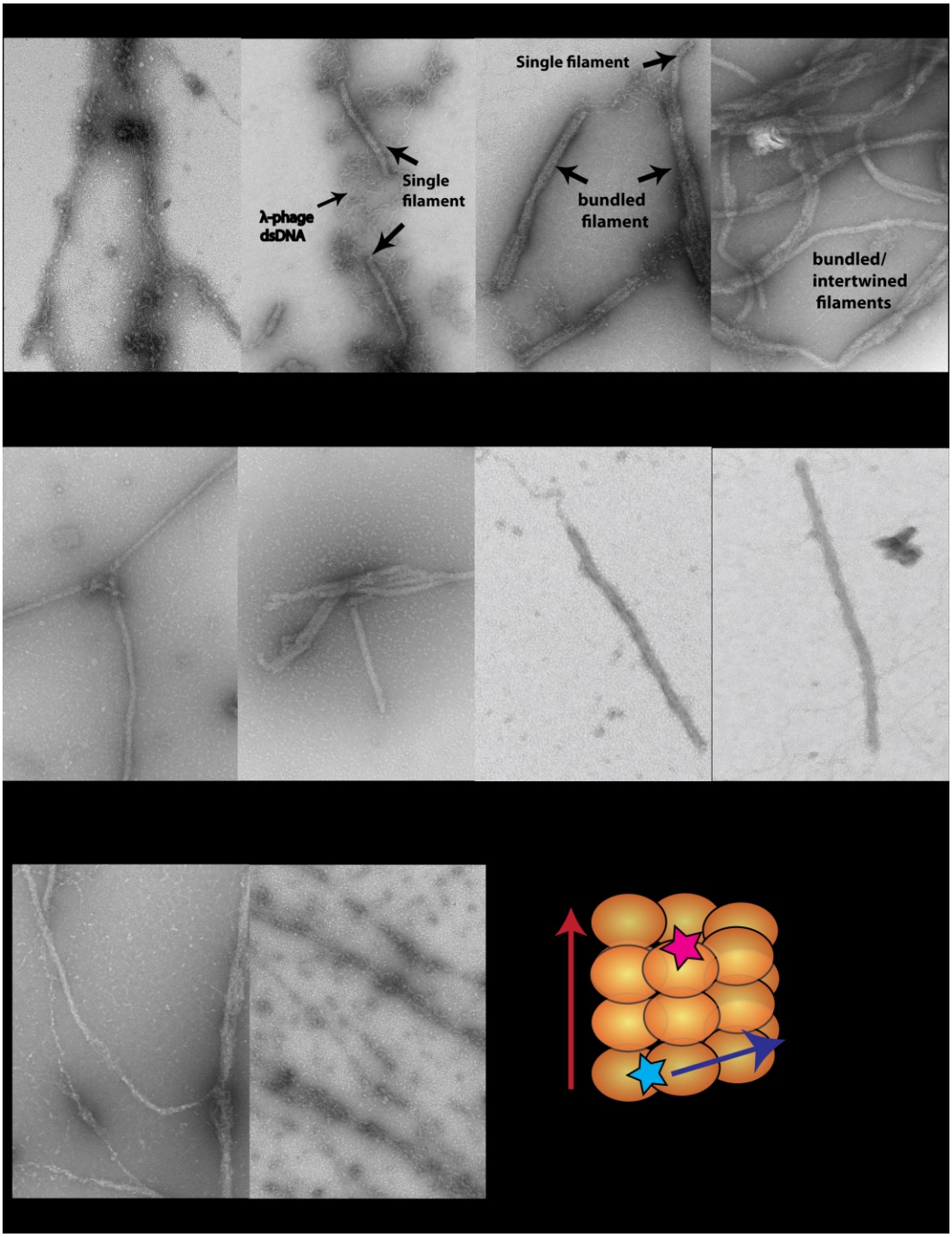

Figure 3.12: AIM2^{PYD} is required to assemble filamentous structures on dsDNA. (A) An electron micrograph of AIM2^{Hin} clusters on λ-phage dsDNA. (B-D) Electron micrographs of wild-type AIM2^{FL} filaments assembled on λ-phage dsDNA**.** (E-J) Electron micrographs of AIM2^{FL} mutants bound to λ-phage dsDNA. (K) A cartoon of the AIM2^{PYD} filament and the locations of mutated side-chains based on the congruent helical symmetry between AIM2^{PYD} and ASC^{PYD} (see also Figure 3.11B).

3.3 Discussion

Presented here is evidence supporting an oligomerization-driven activation mechanism for initiating the assembly of the AIM2 inflammasome in the absence of the currently prevailing autoinhibitory mechanism [31] [32] (**Figure 3.13**). For instance, the basal concentration of AIM2 is presumably low nanomolar under normal conditions, but is dramatically raised on pathogenic invasion (AIM2 is overexpressed by type-1 interferons by at least 200-fold) [19] [20] [38]. The model presented here predicts that, in the absence of cytosolic dsDNA, AIM2 would fail to oligomerize and induce the polymerization of ASC due to its low basal concentration. For instance, in physiologically relevant reaction conditions, AIM2 can assemble into a filament on dsDNA larger than 300 bp even at pico-molar concentrations, but it requires nearly 10,000-fold more AIM2 molecules to auto-assemble filaments without dsDNA (**Figure 3.5** and **Figure 3.8**). On the other hand, because oligomerization is coupled with dsDNA binding, individual AIM2 molecules are capable of finding one another even in the presence of excess dsDNA (**Figure 3.8**). Thus, when foreign dsDNA invades, basal AIM2 would rapidly cluster into a seed filament necessary to nucleate the polymerization of ASC.

The assembly of Rig-I and MDA5 filaments on foreign dsRNA is intrinsically regulated by the ATP turnover at their helicase domains [27] [28] [29]; the filament assembly of Rig-I is further regulated by the recognition of 5′-triphosphate of dsRNA [1] [29]. However, ALRs lack any dsDNA sequence specificity and auto-assemble into filaments without any cofactors [30]. These observations raise the question of whether ALRs are regulated at all. The dsDNA length-dependent binding may

provide a key to answering this question. Although MDA5 can assemble into filaments along the length of dsRNA, the ability to discriminate between 'short' and 'long' dsRNA by MDA5 is not as pronounced as that of AIM2 or IFI16 [27], as if separating the ligand-binding domain from the polymerization domain generates greater dependence on the size of the nucleic acid scaffold for assembly. 'Long,' naked cytosolic dsDNA is rare, and Knipe and colleagues postulated that IFI16 would selectively recognize foreign dsDNA by the degree of chromatinization [12] [39] [40]. Indeed, the highly cooperative relationship between the binding affinity and the size of dsDNA can clearly define an 'off' and 'on' state for assembling the AIM2 nucleoprotein filament (**Figure 3.5** and **Figure 3.8**). Moreover, the size of dsDNA required to build the 'threshold oligomer' (**Figure 3.5** and **Figure 3.8**) also correlates with a previous *in vivo* study in which about 80-bp cytosolic dsDNA was required to induce robust interleukin-1β secretion [31]. Taken together, these data suggest that the size of dsDNA can act as a powerful 'molecular ruler' that can regulate the initiation of the AIM2 inflammasome assembly in a switch-like mechanism.

Unlike the HIN200 domains of IFI16 (see previous chapter), $AIM2^{Hin}$ oligomerizes on dsDNA, which appears to be at least partially responsible for the more robust dsDNA-binding activity of AIM2 than IFI16 *in vitro* [31]. It is tempting to speculate that the oligomerization of $AIM2^{Hin}$ has not been selected against because cytosolic AIM2 is much less likely to encounter self-dsDNA than nuclear IFI16. On the other hand, $p202^{HinB}$ does not bind dsDNA, but forms a tetramer that prevents AIM2 from clustering on dsDNA [37]. The oligomerization activity of

AIM2^{Hin} further strengthens this finding, as p202^{HinB} would physically interfere with clustering of AIM2^{Hin} on dsDNA. Moreover, not all the side chains implicated in the tetramerization of p202^{HinB} are conserved in AIM2 [37] (**Figure 3.7A,B**), thus preventing spurious auto-oligomerization. Nevertheless, the above experiments show that isolated AIM2^{Hin} oligomerizes into random clusters on dsDNA without any filamentous architecture (**Figure 3.12A**) and that disrupting the oligomerization of AIM2^{Hin} still results in filaments isomorphic to wild type (**Figure 3.12E–G**). Thus, the role of AIM2^{Hin} oligomerization is essential, yet limited to generating a 'seed' nucleation unit preceding the filament growth (**Figure 3.13**).

AIM2 is overexpressed by type-1 interferon pathways on pathogenic invasion [30] [38]. The finding of auto-assembly suggest that in principle, dsDNA binding is not *a priori* required to initiate assembly. The auto-oligomerization of AIM2 (**Figure 3.9**) may also enhance the host defense response (**Figure 3.13**). For example, the pre-assembled AIM2 platforms would be able to survey the cytoplasm more effectively because of the larger contiguous surface area. The AIM2^{Hin} clusters could also increase the dsDNA-binding activity via avidity, and the AIM2 oligomers would immediately nucleate the polymerization of ASC without binding foreign dsDNA. However, this could be a double-edged sword: the auto-oligomerization activity of AIM2 could also underlie several autoimmune disorders in which AIM2 is overexpressed via hyperactive interferon pathways [14] [15] [16], as it could cause persistent inflammasome activity without any pathogenic dsDNA.

The symmetry of the AIM2 filament suggests the structure–activity relationships in the assembly. For instance, the three-fold symmetry of the six-start

helix correlates with the dimeric minimal binding unit of AIM2 observed in the FRET assays (**Figure 3.8B,C**; 2 × 3=6). The requirement for clustering about six AIM2 molecules (**Figure 3.5C** and **Figure 3.8C**) to generate a robust 'threshold oligomer' also correlates with one hexameric base of the six-start helix. Furthermore, the optimal oligomer size of 20–25 AIM2 molecules (**Figure 3.5C** and **Figure 3.8C**) also suggests that about four hexameric rings may need to stack up to assemble an optimally stable AIM2 nucleoprotein filament. Finally, because the ASC^{PYD} filament also has the same helical architecture as the $AIM2^{PYD}$ filament, it is tempting to speculate that the assembly of the upstream filament is directly coupled to the downstream effector activation, thus achieving maximal cooperativity.

Both PYDs and CARDs belong to the death-domain (DD) superfamily [35] [41]. Despite their widely variable primary sequences, all known CARDs and PYDs share essentially the same tertiary structure (six-helix bundles) [35] [42]. Several DD proteins are capable of assembling into helical filaments [7] [10] [42], and the PYDs are mostly distinguished from the CARDs by one extended loop region between helices 2 and 3 [35] [41]. Thus, a major outstanding question regarding the signalling mechanism of both PYDs and CARDs has been how one DD protein specifically selects its interacting partner. Although all DD proteins utilize essentially the same set of interaction surfaces for oligomerization [35] [41], it is becoming clear that the resulting oligomers display vastly diverse helical architectures [7] [10] [42]. Indeed, it was recently shown that identical helical symmetry underlies the mechanism by which the Rig-I^{CARD} tetramers nucleate the $MAVS^{CARD}$ filament [10]. It had yet to be tested whether such a symmetric

interaction is a unifying theme in assembling filamentous supramolecular signalling platforms by both CARDs and PYDs, especially when both upstream and downstream DD domains can assemble into infinite filaments as observed from AIM2 and ASC. The consistent helical symmetry between the $AIM2^{PYD}$ and ASC^{PYD} filaments suggests that the corresponding polymerization trajectory between the upstream and downstream oligomers not only underpins the assembly of the inflammasomes but also can be a key to defining the specificity of the DD-family proteins.

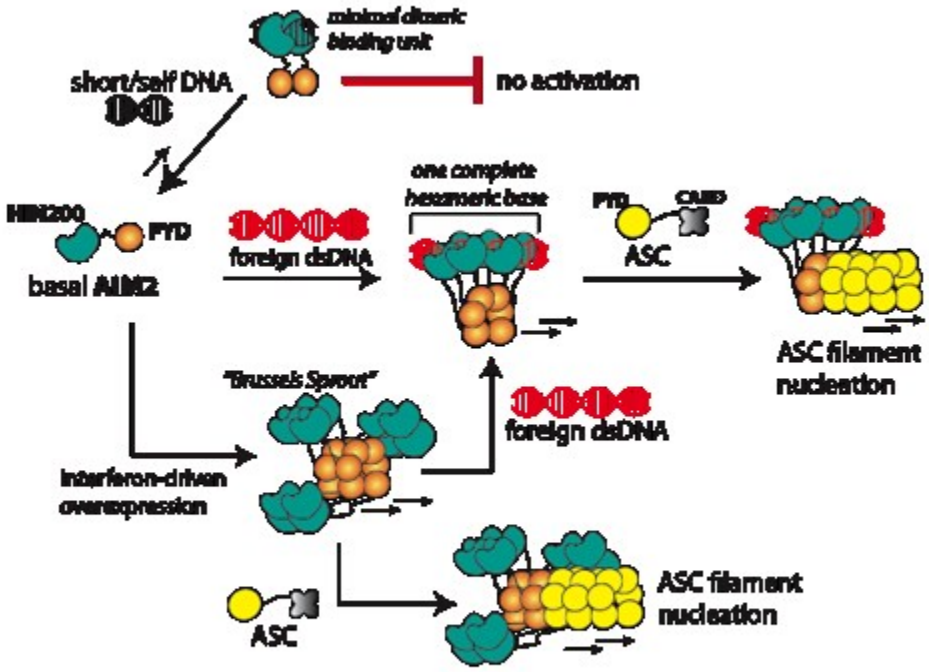

Figure 3.13: A model for the assembly of the AIM2 inflammasome. $AIM2^{PYD}$ is not auto-regulated. Depending on its cellular concentration, auto-assembly or dsDNA-mediated assembly will drive the initial filament formation. Importantly, basal AIM2 requires large dsDNA to generate energetically stable nucleoprotein complexes, because oligomerization is integral to dsDNA binding. The oligomerization of $AIM2^{Hin}$ is important for auto- or dsDNA-mediated filament assembly, but the construction of the filamentous architecture is dictated by $AIM2^{PYD}$. The resulting AIM2 filaments then nucleate the assembly of the ASC filaments via corresponding helical architecture (ASC^{CARD} is not shown in the filament for simplicity).

3.4 Methods

3.4.1 Reagents.

All DNA below 90 bp were purchased from Integrated DNA Technologies, and DNA of greater length was synthesized by PCR. FAM-labelled DNA was also purchased from Integrated DNA Technologies. DyLight-550 and DyLight-650 maleimides were purchased from Thermo.

3.4.2 Recombinant AIM2 constructs.

Full-length AIM2 variants (residues 1–343) were cloned into a pET21 vector (Novagen) with a modified N-terminal MBP tag with a TEV protease recognition site. HIN200 variants (residues 144–343) were cloned into a pET28 vector with a modified N-terminal small ubiquitin-like modifier (SUMO) protein tag or the previously described pET21 vector for assays including an MBP tag. All clones were transformed into *Escherichia coli* strain ER2566 (NEB). Cells were grown at 37 °C to OD_{600} 0.4–0.7, induced with 0.2 mM isopropyl β-D-1-thiogalactopyranoside, expressed for at least 15 h at 18 °C, and then harvested.

For purification of unlabelled protein, cell pellets were resuspended in 20 mM HEPES pH 7.4, 400 mM NaCl, 5% glycerol, 1 mM EDTA and 1 mM dithiothreitol (DTT) while SUMO-tagged construct cell pellets were resuspended in 20 mM HEPES pH 7.4, 400 mM NaCl, 5% glycerol, 20 mM imidazole and 3 mM beta-mercaptoethanolamine (βME); a protease cocktail consisting of phenylmethyl sulfonyl fluoride, benzamidine, leupeptin and pepstatin A was added, as well as lysozyme and DNase I. Cells were lysed by sonication and the insoluble fraction was removed by centrifugation. The supernatants were applied to amylose resin (NEB)

for MBP-tagged constructs or Ni-NTA resin (QIAgen) for SUMO-tagged constructs. A wash of 10–15 CV of resuspension buffer was applied, and then the protein was eluted in either 20 mM HEPES pH 7.4, 120 mM NaCl, 2% glycerol, 30 mM maltose and 3 mM βME for MBP-tagged or 20 mM HEPES pH 7.4, 120 mM NaCl, 2% glycerol, 300 mM imidazole and 3 mM βME for SUMO-tagged constructs. The elution was then applied to a HiTrap-SP Column (GE Healthcare) and the protein eluted off a gradient of 20 mM HEPES pH 7.4, 120 mM–1M NaCl, 2% glycerol and 3 mM βME. Fractions containing highly purified protein were collected. For biochemical assays using MBP-tagged protein, these fractions were then applied to a Superdex200 16/600 gel-filtration column (GE Healthcare) equilibrated in 20 mM HEPES pH 7.4, 400 mM KCl, 2% glycerol, 1 mM EDTA and 1 mM DTT. For biochemical assays using untagged protein, the pooled fractions were diluted 10-fold with 20 mM HEPES pH 7.4, 750 mM NaCl, 2% glycerol, 20 mM imidazole and 3 mM βME, and TEV (MBP tag) or Ulp-1 (SUMO tag) was added, and the solution was dialysed at least 15 h against 20 mM HEPES pH 7.4, 120 mM NaCl, 2% glycerol, 20 mM imidazole and 3 mM βME. For full length, the solution was then re-applied to the HiTrap-SP column, and then eluted off in tandem with a HisTrap-FF (GE Healthcare) with 20 mM HEPES pH 7.4, 750 mM NaCl, 2% glycerol, 20 mM imidazole and 3 mM βME. For HIN200, the solution was applied to a HisTrap-FF column in tandem with a HiTrap-SP column, the HisTrap column was then removed and the protein was eluted from the HiTrap column in the same buffer as the full length. The elution fractions were applied directly to a Superdex75 16/600 gel-filtration column (GE Healthcare) equilibrated

in 20 mM HEPES pH 7.4, 400 mM KCl, 2% glycerol, 1 mM EDTA and 1 mM DTT. Fractions containing the protein were then concentrated.

For fluorescent labelling of proteins, after elution from their respective affinity columns, they were applied to a HiTrap-SP column and eluted using a gradient of 20 mM HEPES pH 7.4, 120 mM NaCl–1M NaCl, 2% glycerol and 1 mM TCEP. The solution was then divided into two, and to one was added DyLight-550 at a concentration of 4 × (AIM2) and to the other was added DyLight-650 at a concentration of 4 × (AIM2). Labelling proceeded for at least 15 h. Excess dye was quenched using 10 mM βME, and the solution was re-applied to their respective affinity columns. After extensive washing with 20 mM HEPES pH 7.4, 750 mM NaCl, 2% glycerol and 3 mM βME until no fluorescence could be detected in the flow through, the protein was eluted in wash buffer supplemented with either 30 mM maltose (full length) or 300 mM imidazole (HIN200). Cleavage and further purification proceeded as above. The dye:protein ratio was then calculated as per the manufacturer's instructions, which was ~1:1.

3.4.3 Biochemical assays.

All experiments were performed at least three times, the fits to data were generated by Kaleidagraph software (Synergy). Fluorescence anisotropy-binding experiments were carried out in either 40 mM HEPES pH 7.4, 160 mM KCl, 5% glycerol, 0.1% Triton X-100, 1 mM EDTA, 5 mM DTT (herein referred to as 'buffer A') or 40 mM HEPES pH 7.4, 400 mM KCl, 5% glycerol, 0.1% Triton X-100, 1 mM EDTA and 5 mM DTT (herein referred to as 'buffer B') at room temperature as described in the previous chapter.

Förster resonance energy transfer experiments were carried out in either buffer A (for $AIM2^{FL}$ and $AIM2^{Hin}$) or buffer C (40 mM HEPES pH 7.4, 60 mM KCl, 5% glycerol, 0.1% Triton X-100, 1 mM EDTA and 5 mM DTT) for $AIM2^{Hin}$ as described in the previous chapter.

Electrophoretic mobility shift assay experiments were carried out in buffer A. To a fixed amount of fluorescein-labelled dsVACV72 was added increasing concentrations of AIM2. The reaction was allowed to equilibrate at room temperature (at least 20 min), then applied to a 4% 116:1 acrylamide:bis-acrylamide Tris-borate-EDTA gel. The gel was run at 100 V in 1 × Tris-borate-EDTA buffer and imaged using a Typhoon imager (GE Healthcare; excitation at 488 nm, emission at 532 nm).

3.4.4 Electron microscopy.

AIM2 samples were adsorbed to glow-discharged carbon grids for 2 min, then blotted and transferred through two consecutive drops of 1% uranyl formate or 1% uranyl acetate for a total of 1–2 min. The carbon film was then quickly dried by aspiration. Images were collected with either a Philips BioTwin CM120 (FEI) at Johns Hopkins School of Medicine or Tecnai 12 at University of Virginia. For the dsDNA·AIM2 complexes, λdsDNA and AIM2 constructs were incubated 30 min before EM sample preparation. The linearized plasmid was generated by digesting the pET28b vector (Novagen) with BamH1 (NEB). The nicked circle was generated by cutting pET28b with Nt.BspQI (NEB). The modified vectors were then agarose-gel purified.

3.4.5 Homology modeling.

The crystal structure of $AIM2^{PYD}$ (PDB ID: 3VD8) was aligned to individual protomers of the cryo-EM structure of the ASC^{PYD} filament (PDB ID: 3J63) using Pymol (The PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger, LLC). The root mean squared deviation of the individual alignment is <1 Å.

3.4.6 Symmetry determination.

Micrographs of negatively stained $AIM2^{FL}$ filaments were scanned using a Nikon CoolPix 8000 with a raster of 4.16 Å/pixel. Filaments were extracted using the e2helixboxer routine within EMAN2 [43], and the SPIDER software package [44] was used for subsequent steps. Overlapping boxes 96 pixels long were cut from these filaments, and 7,607 boxes were aligned against a preliminary reconstruction and windowed to 30 pixels to generate an averaged power spectrum.

3.5 References

[1] M Yoneyama, K Onomoto, M Jogi, T Akaboshi, and T. Fujita, "Viral RNA detection by RIG-I-like receptors," *Curr Opin Immunol*, vol. 32, pp. 48-53, Feb 2015.

[2] Y Del Toro Duany, B Wu, and S Hur, "MDA5-filament, dynamics and disease," *Curr. Opin. Virol.*, vol. 12C, pp. 20-25, 2015.

[3] A Lu and H Wu, "Structural mechanisms of inflammasome assembly," *FEBS J*, vol. 282, pp. 435-444, 2015.

[4] D.M. Knipe, "Nuclear sensing of viral DNA, epigenetic regulation of herpes simplex virus infection, and innate immunity," *Virology*, vol. 479-480C, pp. 153-159, 2015.

[5] H Wu, "Higher-order assemblies in a new paradigm of signal transduction," *Cell*, vol. 153, pp. 287-292, 2013.

[6] JC, Magupalli, VG, Wu, H Kagan, "SMOCs: supramolecular organizing centres that control innate immunity," *Nat Rev Immunol*, vol. 14, pp. 821-826, 2014.

[7] A Lu et al., "Unified polymerization mechanism for the assembly of ASC-dependent inflammasomes," *Cell*, vol. 156, no. 6, pp. 1193-1206, Mar 2014.

[8] X Cai et al., "Prion-like polymerization underlies signal transduction in antiviral immune defense and inflammasome activation," *Cell*, vol. 156, no. 6, pp. 1207-22, Mar 2014.

[9] B Wu et al., "Molecular imprinting as a signal-activation mechanism of the viral RNA sensor RIG-I," *Mol Cell*, vol. 55, no. 4, pp. 511-23, Aug 2014.

[10] B Wu et al., "Structural basis for dsRNA recognition, filament formation, and antiviral signal activation by MDA5," *Cell*, vol. 152, no. 1-2, pp. 276-89, Jan 2013.

[11] T Fernandes-Alnemri et al., "The AIM2 inflammasome is critical for innate immunity to Francisella tularensis," *Nat Immunol*, vol. 11, no. 5, pp. 385-93, May 2010.

[12] MH Orzalli, SE Conwell, C Berrios, JA DeCaprio, and DM Knipe, "Nuclear interferon-inducible protein 16 promotes silencing of herpesviral and transfected DNA," *Proc Natl Acad Sci U S A*, vol. 110, no. 47, pp. E4492-501, Nov 2013.

[13] Q Feng et al., "MDA5 detects the double-stranded RNA replicative form in picornavirus-infected cells," *Cell Rep*, vol. 2, no. 5, pp. 1187-96, Nov 2012.

[14] M Mondini et al., "Role of the interferon-inducible gene IFI16 in the etiopathogenesis of systemic autoimmune disorders," *Ann N Y Acad Sci*, vol. 1110, pp. 47-56, Sept 2007.

[15] W Zhang et al., "AIM2 facilitates the apoptotic DNA-induced systemic lupus erythematosus via arbitrating macrophage functional maturation," *J Clin Immunol*, vol. 33, no. 5, pp. 925-37, Jul 2013.

[16] S Smith and C Jefferies, "Role of DNA/RNA sensors and contribution to autoimmunity," *Cytokine Growth Factor Rev*, vol. 25, no. 6, pp. 745-57, Dec

2014.

[17] V Hornung et al., "AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC," *Nature*, vol. 458, no. 7237, pp. 514-8, Mar 2009.

[18] T Burckstummer et al., "An orthogonal proteomic-genomic screen identifies AIM2 as a cytoplasmic DNA sensor for the inflammasome," *Nat Immunol*, vol. 10, no. 3, pp. 266-72, Mar 2009.

[19] TL Roberts et al., "HIN-200 proteins regulate caspase activation in response to foreign cytoplasmic DNA," *Science*, vol. 323, no. 5917, pp. 1057-60, Feb 2009.

[20] T Fernandes-Alnemri, JW Yu, P Datta, J Wu, and ES Alnemri, "AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA," *Nature*, vol. 458, no. 7237, pp. 509-13, Mar 2009.

[21] L Unterholzner et al., "IFI16 is an innate immune sensor for intracellular DNA," *Nat Immunol*, vol. 11, no. 11, pp. 997-1004, Nov 2010.

[22] T Li, J Chen, and IM Cristea, "Human cytomegalovirus tegument protein pUL83 inhibits IFI16-mediated DNA sensing for immune evasion," *Cell Host Microbe*, vol. 14, no. 5, pp. 591-9, Nov 2013.

[23] N Kerur et al., "IFI16 acts as a nuclear pathogen sensor to induce the inflammasome in response to Kaposi Sarcoma-associated herpesvirus infection," *Cell Host Microbe*, vol. 9, no. 5, pp. 363-75, May 2011.

[24] KM Monroe et al., "IFI16 DNA sensor is required for death of lymphoid CD4 T cells abortively infected with HIV," *Science*, vol. 343, no. 6169, pp. 428-32, Jan

2014.

[25] MS VanLoock et al., "ATP-mediated conformational changes in the RecA filament," *Structure*, vol. 11, no. 2, pp. 187-96, Feb 2003.

[26] VE Galkin et al., "The Rad51/RadA N-terminal domain activates nucleoprotein filament ATPase activity," *Structure*, vol. 14, no. 6, pp. 983-92, Jun 2006.

[27] A Peisley et al., "Kinetic mechanism for viral dsRNA length discrimination by MDA5 filaments," *Proc Natl Acad Sci U S A*, vol. 109, no. 49, pp. E3340-9, Dec 2012.

[28] A Peisley et al., "Cooperative assembly and dynamic disassembly of MDA5 filaments for viral dsRNA recognition," *Proc Natl Acad Sci U S A*, vol. 108, no. 52, pp. 21010-5, Dec 2011.

[29] A Peisley, B Wu, H Yao, T Walz, and S Hur, "RIG-I forms signaling-competent filaments in an ATP-dependent, ubiquitin-independent manner," *Mol Cell*, vol. 51, no. 5, pp. 573-83, Sep 2013.

[30] SA Schattgen and KA Fitzgerald, "The PYHIN protein family as mediators of host defenses," *Immunol Rev*, vol. 242, no. 1, pp. 109-18, Sep 2011.

[31] T Jin et al., "Structures of the HIN domain:DNA complexes reveal ligand binding and activation mechanisms of the AIM2 inflammasome and IFI16 receptor," *Immunity*, vol. 36, no. 4, pp. 561-71, Apr 2012.

[32] T Jin, A Perry, P Smith, J Jiang, and TS Xiao, "Structure of the absent in melanoma 2 (AIM2) pyrin domain provides insights into the mechanisms of AIM2 autoinhibition and inflammasome assembly," *J Biol Chem*, vol. 288, no. 19,

pp. 13225-35, May 2013.

[33] A Lu, V Kabaleeswaran, T Fu, VG Magupalli, and H Wu, "Crystal structure of the F27G AIM2 PYD mutant and similarities of its self-association to DED/DED interactions," *J Mol Biol*, vol. 426, no. 7, pp. 1420-7, Apr 2014.

[34] PR Vajjhala, RE Mirams, and JM Hill, "Multiple binding sites on the pyrin domain of ASC protein allow self-association and interaction with NLRP3 protein," *J Biol Chem*, vol. 287, no. 50, pp. 41732-43, Dec 2012.

[35] RA Ratsimandresey, A Dorfleutner, and C Stehlik, "An Update on PYRIN Domain-Containing Pattern Recognition Receptors: From Immunity to Pathology," *Front Immunol*, vol. 4, p. 440, Dec 2013.

[36] E Latz, TS Xiao, and A Stutz, "Activation and regulation of the inflammasomes," *Nat Rev Immunol*, vol. 13, no. 6, pp. 397-411, Jun 2013.

[37] Q Yin et al., "Molecular mechanism for p202-mediated specific inhibition of AIM2 inflammasome activation," *Cell Rep*, vol. 4, no. 2, pp. 327-39, Jul 2013.

[38] S Veeranki, X Duan, R Panchanathan, H Liu, and D Choubey, "IFI16 protein mediates the anti-inflammatory actions of the type-I interferons through suppression of activation of caspase-1 by inflammasomes," *PLoS ONE*, vol. 6, no. 10, p. e27040, Oct 2011.

[39] MH Orzalli and DH Knipe, "Cellular sensing of viral DNA and viral evasion mechanisms," *Annu Rev Microbiol*, vol. 68, pp. 477-92, Jun 2014.

[40] MH Orzalli, NA DeLuca, and DM Knipe, "Nuclear IFI16 induction of IRF-3 signaling during herpesviral infection and degradation of IFI16 by the viral

ICP0 protein," *Proc Natl Acad Sci U S A*, vol. 109, no. 44, pp. E3008-17, Oct 2012.

[41] HH Park et al., "The death domain superfamily in intracellular signaling of apoptosis and inflammation," *Annu Rev Immunol*, vol. 25, pp. 561-86, 2007.

[42] Q Qiao et al., "Structural architecture of the CARMA1/Bcl10/MALT1 signalosome: nucleation-induced filamentous assembly," *Mol Cell*, vol. 51, no. 6, pp. 766-79, Sep 2013.

[43] G Tang et al., "EMAN2: an extensible image processing suite for electron microscopy," *J Struct Biol*, vol. 157, no. 1, pp. 38-46, Jan 2007.

[44] J Frank et al., "SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields," *J Struct Biol*, vol. 116, no. 1, pp. 190-9, Jan-Feb 1996.

Chapter 4

A Mechanism for Discriminating Self from Nonself for IFI16

Reproduced in part from:

Stratmann S, Morrone S, van Oijen AM, Sohn J. (2015). The innate immune sensor IFI16 recognizes foreign DNA in the nucleus by scanning along the duplex. *Elife*. Dec 16;4. pii: e11721.

4.1 Introduction

The host innate immune system detects infection by directly recognizing molecular signatures associated with pathogens [1] [2]. Remarkably, such signatures include universal building blocks of all life, such as DNA and RNA [3] [4] [5]. In the cytoplasm, the immune system relies on the absence of endogenous DNA, and thus marks all detected DNA as "foreign" (nonself) [4] [5]. However, DNA viruses often evade the cytosolic detection machineries, as their genomes are not exposed until reaching the nucleus [4] [5]. The host counters this infection strategy in the nucleus by directly assembling supramolecular signaling platforms that trigger inflammatory responses on invading foreign DNA, but not on its own genomic material [6] [7] [8]. Although key players that target foreign dsDNA in the host nucleus have been identified [4] [5], the molecular mechanisms by which these sensors distinguish self from nonself dsDNA remain unknown.

The interferon-inducible protein 16 (IFI16) is a key innate immune sensor that detects foreign dsDNA and uses it as a scaffold to assemble supra-molecular signaling platforms in both the host nucleus and cytoplasm [6] [7] [8] [9] [10] (**Figure 4.1**). IFI16 plays a central role in defense against a number of pathogens (e.g., herpes simplex virus-1) [6] [7] [8] [9] [10] ; on the other hand, persistent IFI16 signaling is associated with autoimmunity (e.g. Sjögren's syndrome) [11] [12] [13] [14] [15]. The molecular mechanisms by which IFI16 selectively targets foreign dsDNA remain unknown. To establish a functional signaling platform, IFI16 must overcome two challenges. First, individual IFI16 molecules must be able to locate one another on large pathogen genomes with sizes ranging from 10^5 to 10^6 base

pairs (bps). Second and more importantly, this assembly mechanism can only take place on foreign dsDNA and must be inhibited on host dsDNA (**Figure 4.1**). Here, a unifying molecular mechanism that explains how IFI16 resolves these central issues in initiating its foreign-dsDNA sensing pathways is investigated.

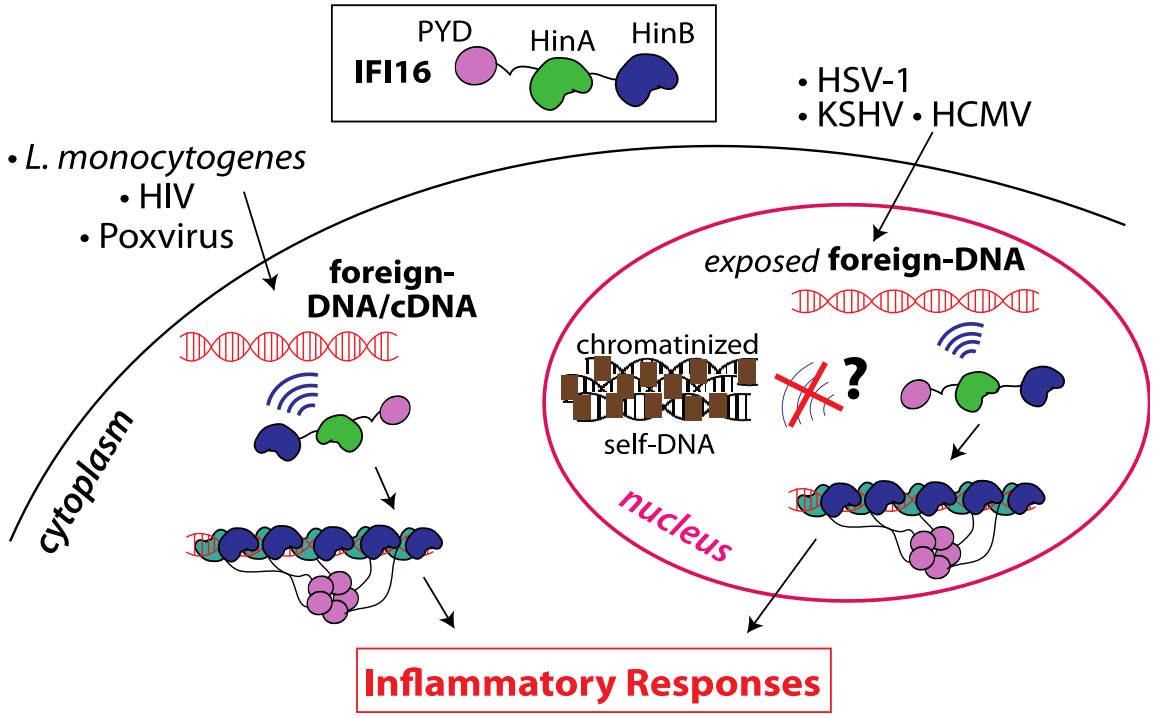

Figure 4.1: A simplified picture of the path of IFI16. Top: IFI16 is composed of three functional domains, namely one pyrin domain (PYD) and two dsDNA-binding Hin domains (HinA and HinB; Hin: hematopoietic interferon inducible nuclear antigen). flanked by unstructured regions. Bottom: IFI16 is found in both the cytoplasm and nucleus, where it exhibits differentiated responses to viruses depending upon its location.

4.2 Results

4.2.1. IFI16 Length-Dependent Assembly Kinetics Suggests 1D Diffusion

To identify the mechanisms underlying assembly of IFI16 signaling platforms on DNA, the oligomerization kinetics of FRET donor and acceptor labeled IFI16 on naked dsDNA (FRET: Förster resonance energy transfer; **Figure 4.2** and **Table 4.1**) was monitored. Chapter 2 demonstrated the existence of such oligomers and reported on their equilibrium binding properties, but did not provide insights into the assembly mechanisms. Using various dsDNA fragment sizes present in excess, the assembly rate increased non-linearly and by 50-fold from 60 to 200 bps dsDNA, above which it stayed constant (up to 600 bps; **Figure 4.2A**, **Figure 4.2B**). With a dsDNA-binding footprint of ~15 bp for one IFI16 (see Chapter 2), these results indicate that about four copies are required to initiate assembly, and about ten IFI16 molecules are required for optimal oligomeric assembly (**Figure 4.2B**). Further, the assembly rate constants scaled linearly with the IFI16 concentration for all measured DNA lengths (**Table 4.1**), indicating that a purely cooperative assembly mechanism is unlikely. In line with this, Chapter 2 showed relatively small contributions of cooperativity in oligomerization with Hill constants near 2 for DNA substrates up to 2000 bp.

These observations of the faster assembly on longer dsDNA suggest a model in which IFI16 scans along dsDNA to increase the probability of encountering other IFI16 molecules (**Figure 4.3**). The 1D diffusion of IFI16 on dsDNA explains why the assembly rates increase with the DNA length in the bulk experiments (**Figure 4.2B**). With the longer dsDNA acting as an antenna, it allows binding of more IFI16 while

1D diffusion facilitates dynamic association (**Figure 4.2B**). The saturation of the assembly rate (**Figure 4.2B**) can be explained by the square dependence of the diffusional search time on length: at a sufficiently long dsDNA length, the dissociation rate of an individual IFI16 will be faster than the time needed to scan along the entire length of the DNA. Overall, the results of these experiments are consistent with the dsDNA-size dependent binding *in vitro* (see Chapter 2), which also correlates with the IFI16-induced inflammatory responses *in vivo* [9]. Thus, it is proposed that the 1D-diffusion mediated assembly plays a key role in regulating the overall IFI16-mediated immune responses.

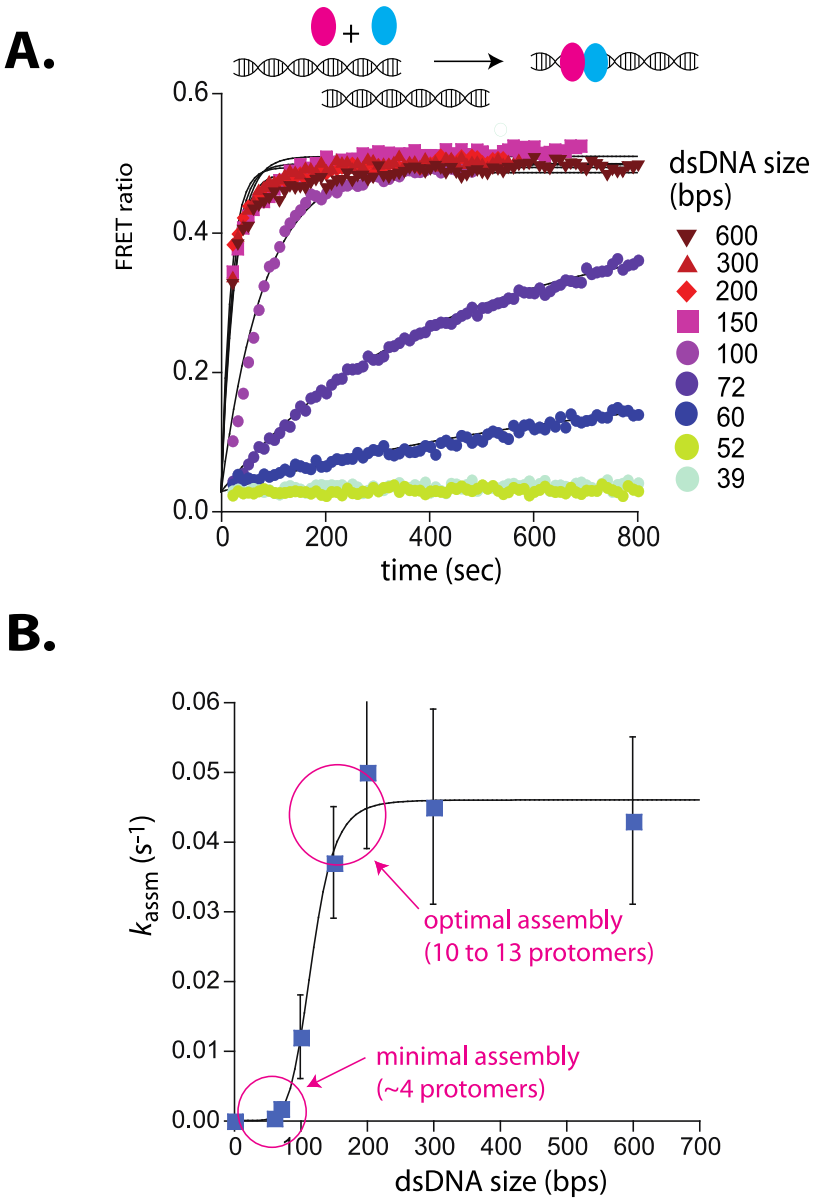

Figure 4.2: IFI16 assembles faster on longer dsDNA. (A) Top: a cartoon scheme for FRET experiments. The two differentially colored ovals represent fluorescently (DyLight-550 and DyLight-650) labeled IFI16. Bottom: The time-dependent changes in the emission ratio between FRET donor and acceptor labeled IFI16 (50 nM) were monitored at 33 μg/mL of each dsDNA (e.g., 6-fold higher than the dissociation constant for 39-bp dsDNA [see Chapter 2]). Lines are fits to a first-order exponential equation. All shown representative experiments were performed at least three times. (B) A plot of observed assembly rates (k_{assm}) vs. dsDNA sizes (see also Table 4.1).

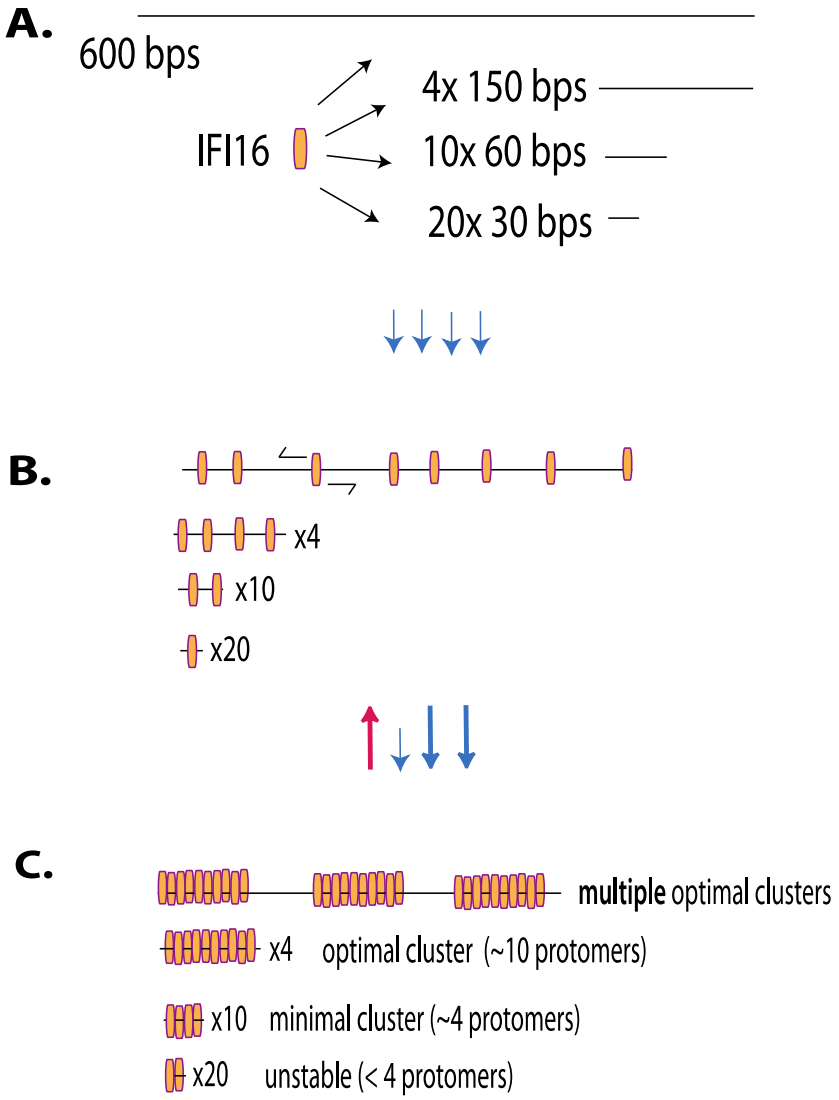

Figure 4.3: A 1D-diffusion assisted assembly mechanism can explain the observed assembly profile of IFI16. (A) At the same mass-concentrations the number of individual dsDNA fragments present in each assay is inversely proportional to the length of dsDNA. (B) Individual IFI16 molecules initially bind dsDNA at random positions and diffuse one-dimensionally while searching for other respective protomers; the number of IFI16 molecules residing on the same dsDNA fragment should be proportional to the length of dsDN (e.g., there are four times more individual 150-bp fragments than 600-bp fragments). (C) IFI16 fails to assemble into an oligomer on dsDNA shorter than 60 bp (indicated by a red arrow pointing up). The saturating rates can be explained if the final FRET signals arise from formation of distinct optimal oligomers.

dsDNA size (bps)	25 nM IFI16 (sec^{-1})	50 nM IFI16 (sec^{-1})
60	0.0004 ± 0.0002	0.0009 ± 0.0002
70	0.0008 ± 0.0002	0.0021 ± 0.0008
100	0.0052 ± 0.0008	0.012 ± 0.005
150	0.019 ± 0.004	0.037 ± 0.008
200	0.021 ± 0.007	0.049 ± 0.011
300	0.029 ± 0.008	0.045 ± 0.014
600	0.021 ± 0.008	0.043 ± 0.012

Table 4.1: dsDNA-mediated oligomerization rates of FRET-labeled IFI16. Each experiment was performed at least three times and errors were calculated by using the standard deviations.

4.2.2. Chromatinization of DNA Inhibits IFI16 Cluster Formation

It has long been speculated that chromatinization acts as the key feature that allows IFI16 to distinguish host from foreign DNA in the nucleus [6] [7] [8] [10] [16] [17]; IFI16 oligomerizes on exposed invading foreign-dsDNA before it becomes hetero-chromatinized. A previous *in vivo* work has demonstrated that pre-chromatinized invading SV40 DNA is able to evade IFI16 oligomerization and downstream responses [16]. Nevertheless, the molecular mechanism by which IFI16 could use chromatinization to distinguish self from nonself has yet to be identified. To directly address this issue, a competition-binding assay was used to investigate how IFI16 interacts with dsDNA fragments containing two nucleosomes with varying spacer sizes (6, 30, 50, and 70 bps; **Figure 4.4A**). Here, di-nucleosomes with 6-, 30-, and 50-bp spacer failed to compete against IFI16-bound FAM-labeled 70-bp dsDNA, (**Figure 4.4B**). On the other hand, the di-nucleosome with 70-bp spacer competed similarly as 70-bp naked dsDNA, but significantly more weakly than naked 300-bp dsDNA (Figure 3B). In FRET assembly assays, di-nucleosomes with spacers shorter than 70-bp failed to support assembly (Figure 3C), consistent with the FRET kinetics assays using naked dsDNA (**Figure 4.2A,B**). The 70-bp spacer di-nucleosome supported oligomerization of IFI16; however, the assembly kinetics was again similar to that of naked 70-bp dsDNA, but not that of naked 300-bp dsDNA (**Figure 4.4C**). Taken together, these results show that efficient IFI16 cluster formation requires a minimal length of 50-70 base pairs of exposed dsDNA. Considering that the size of dsDNA linker between two nucleosomes is about 20 to 30 bps in mammals [18], these results directly support the hypothesis that

chromatinization is a key deterrent for preventing the assembly of IFI16 signaling platforms on self-dsDNA.

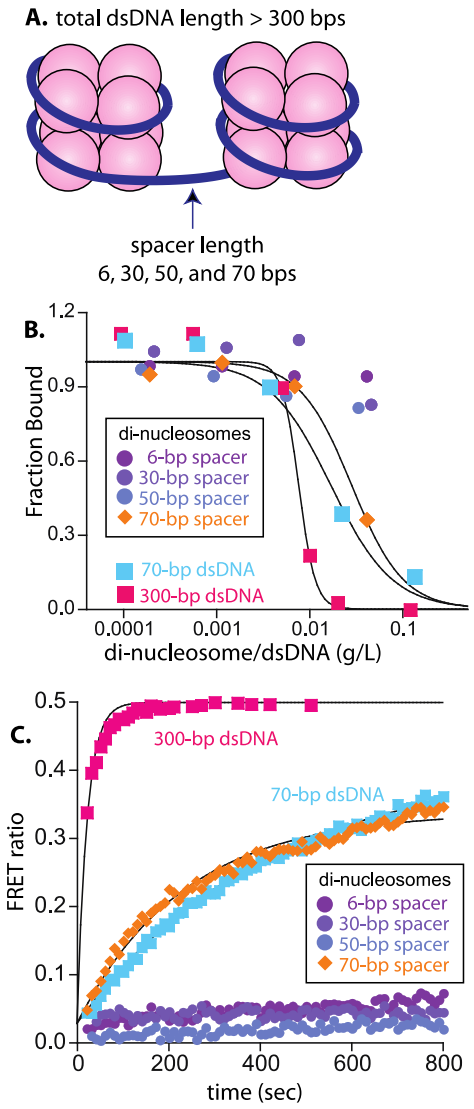

Figure 4.4: Nucleosomes inhibit oligomerization. (A) A cartoon of di-nucleosome constructs with varying dsDNA spacers. (B) Competition binding assays using IFI16 bound FAM-lebeled 70-bp dsDNA against various di-nucleosomes and naked dsDNA. The lines are fits to: 1/(1+([$DNA_{competitor}$]/IC_{50})^Hill Coefficient), where IC50 indicates the concentration of competitor at 50% efficiency. The mass-concentration of each competitor was calculated using dsDNA, but not histones. (C) The time-dependent changes in the emission ratio between the FRET donor and acceptor labeled IFI16 (50 nM) were monitored at 33 μg/mL of each nucleosome or naked dsDNA. The lines are fits to a first-order exponential equation.

4.3 Discussion

The molecular mechanism by which innate immune sensors distinguish self from foreign dsDNA in the host nucleus has been a major unresolved question in innate immunology [6] [7] [8] [10] [16] [17] [19]. The oligomerization of IFI16 on under-chromatinized foreign DNA plays a key role not only in initiating inflammatory and antiviral responses [8] [10] [20],, but also in regulating the hetero-chromatinization and silencing of viral dsDNA [16] [21]. By using time-resolved bulk fluorescence assays, it is demonstrated here that IFI16 scans one-dimensionally along exposed dsDNA to assemble, and that chromatinization is sufficient to inhibit IFI16 from targeting host dsDNA for assembly. While this clustering on dsDNA presents a tempting explanation for IFI16's role in viral gene silencing, future *in vivo* experiments await to test this. IFI16 belongs to the family of AIM2-like receptors, which include other nuclear and cytosolic foreign dsDNA-sensors. It will be interesting to determine whether and how these other related sensors use exposed dsDNA as a 1D "digital ruler" to regulate their signaling platform assembly. This family of sensors is implicated in a number of autoimmune disorders [11] [12] [13] [14] [15]; how regulation of assembly is disrupted may provide insights into these afflictions.

4.4 Methods

4.4.1 Protein Expression and Purification

Full-length human IFI16 was cloned and expressed using *E. Coli* T7 express cells (NEB) as a C-terminally His6-tagged protein as described in Chapter 2.

4.4.2 DNA Ligand Preparation

dsDNA shorter than 90-bp were obtained from Integrated DNA Technologies (IDT) and are the same ligands as listed in Chapter 2. The complementary strands were dissolved and mixed in 1:1 molar ratio, melted at 95°C for 10min, and the temperature was lowered to 25°C at a rate of 1°C/min. Ligands of greater length were obtained by polymerase-chain reaction (PCR) using the Maltose Binding Protein fusion tag cloning sequence as template and primers of appropriate sequence for a final length as indicated in the assays. Plasmids containing the Widom-601/603 sequence with indicated linker lengths were a kind gift of Dr. Gregory Bowman. The nucleosomal DNA was obtained by PCR from these constructs with appropriate primers. All ligands were gel-purified.

4.4.3 Fluorescent Labeling

DyLight-550 or DyLight-650, was incorporated to IFI16 using maleimide chemistry (purchased from Thermo Scientific and Invitrogen) and was performed as described in Chapter 2. Fluorescein-labeled dsDNA72 was obtained from IDT.

4.4.4 Octamer Refolding and Nucleosome Reconstitution

Lyophylized *Xenopus laevis* histones H1A, H2A, H3, and H4 were a kind gift of Dr. Cynthia Wolberger. Octamer refolding and nucleosome reconstitution was

performed as described in Luger et al. [22], at a 2:1 molar ratio of octamer:DNA. An agarose gel of reconstituted nucleosomes is shown in **Figure 4.5**.

4.4.5 Biochemical Assays

All absorption, fluorescence anisotropy, and fluorescence excitation/emission experiments were performed in a Tecan Infinite M1000. All experiments were performed at least three times and the fits to data were generated by Kaleidagraph software (synergy).

4.4.6 Competition Binding Assays

All reactions were performed in 40 mM HEPES pH 7.4, 160 mM KCl, 5% glycerol, 1 mM EDTA, 0.1% triton-X-100, 5 mM DTT (Reaction Buffer). 300 nM IFI16 and 4.5 nM fluorescein-labeled dsVACV72 were incubated together at room temperature for 20min. Increasing concentrations of competing DNA were added to the reaction to a final concentration of 100nM IFI16 and 1.5nM dsVACV72, and the changes in fluorescence anisotropy were recorded as indicated in Chapter 2.

4.4.7 FRET Time Dependence Assays

All reactions were performed in Reaction Buffer. 66 μg/ml of each dsDNA or di-nucleosomes was placed in the plate wells, and the reaction was initiated by adding an equivalent volume of IFI16-550 and IFI16-650 (1:1 molar ratio) to the indicated final concentration. The dead time between addition of IFI16 and the first measurement was 15-20s. The final dsDNA molar-concentrations are at least 6-fold higher than their determined binding constants by fluorescence anisotropy assays described in Chapter 2, and the FRET ratio for each time point was calculated by dividing the acceptor emission (678 nm) by the donor emission (574 nm).

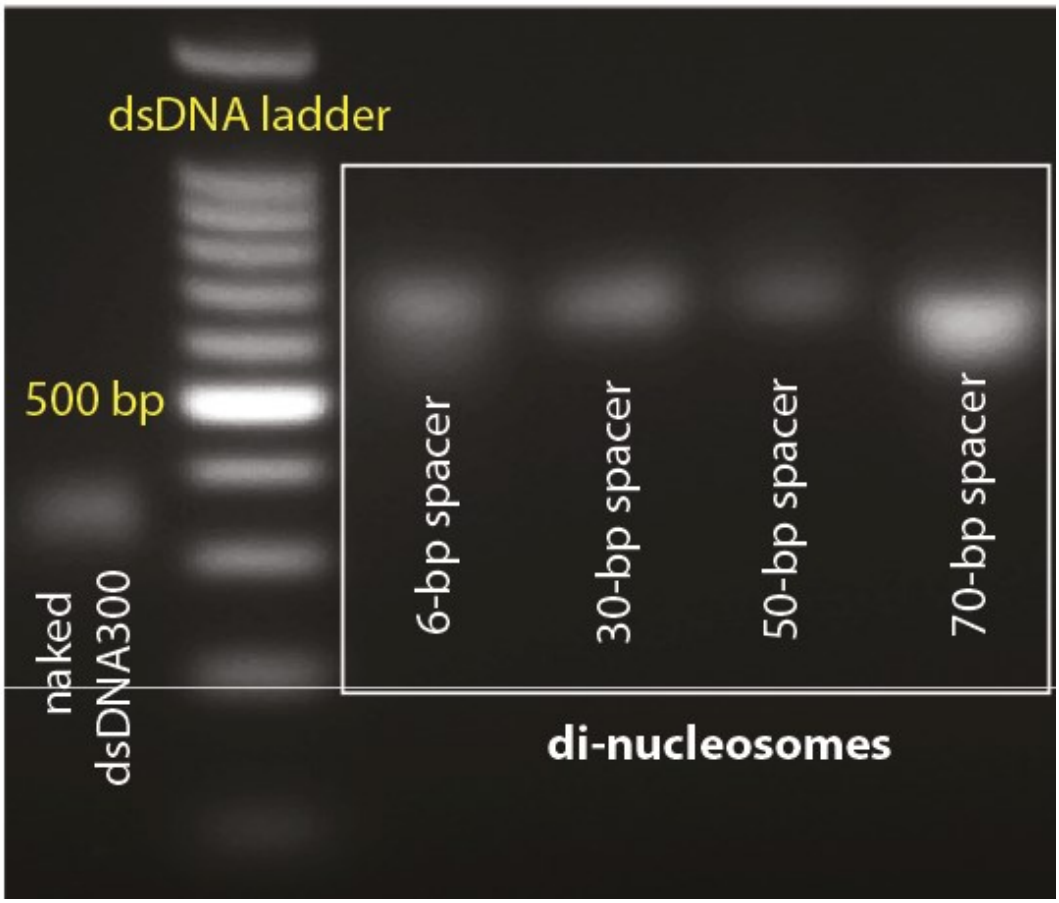

Figure 4.5: Agarose gels with nucleosome preparations. Purified di-nucleosomes for the bulk FRET assays.

4.5 References

[1] Jr CA Janeway and Medzhitov R, "Innate immune recognition," *Annu Rev Immunol*, vol. 20, pp. 197-216, 2002.

[2] R Medzhitov and C, Jr Janeway, "Innate immune recognition: mechanisms and pathways," *Immunol Rev*, vol. 173, pp. 89-97, 2000.

[3] R Medzhitov and CA, Jr Janeway, "Decoding the patterns of self and nonself by the innate immune system," *Science*, vol. 296, pp. 298-300, 2002.

[4] MH Orzalli and DM Knipe, "Cellular sensing of viral DNA and viral evasion mechanisms," *Annu Rev Microbiol*, vol. 68, pp. 477-92, 2014.

[5] SR Paludan and AG Bowie, "Immune sensing of DNA," *Immunity*, vol. 38, pp. 870-80, 2013.

[6] KE Johnson, L Chikoti, and B Chandran, "Herpes simplex virus 1 infection induces activation and subsequent inhibition of the IFI16 and NLRP3 inflammasomes," *J Virol*, vol. 87, pp. 5005-18, 2013.

[7] T Li, BA Diner, J Chen, and IM Cristea, "Acetylation modulates cellular distribution and DNA sensing ability of interferon-inducible protein IFI16," *Proc Natl Acad Sci U S A*, vol. 109, pp. 10558-63, 2012.

[8] N Kerur et al., "IFI16 acts as a nuclear pathogen sensor to induce the inflammasome in response to Kaposi Sarcoma-associated herpesvirus infection," *Cell Host Microbe*, vol. 9, pp. 363-75, 2011.

[9] L Unterholzner et al., "IFI16 is an innate immune sensor for intracellular DNA," *Nat Immunol*, vol. 11, pp. 997-1004, 2010.

[10] MH Orzalli, NA Deluca, and DM Knipe, "Nuclear IFI16 induction of IRF-3 signaling during herpesviral infection and degradation of IFI16 by the viral ICP0 protein," *Proc Natl Acad Sci U S A*, vol. 109, pp. E3008-17, 2012.

[11] M Mondini et al., "Role of the interferon-inducible gene IFI16 in the etiopathogenesis of systemic autoimmune disorders," *Ann N Y Acad Sci*, vol. 1110, pp. 47-56, 2007.

[12] M Mondini et al., "The interferon-inducible HIN-200 gene family in apoptosis and inflammation: implication for autoimmunity," *Autoimmunity*, vol. 43, pp. 226-31, 2010.

[13] D Choubey et al., "Interferon-inducible p200-family proteins as novel sensors of cytoplasmic DNA: role in inflammation and autoimmunity," *J Interferon Cytokine Res*, vol. 30, pp. 371-80, 2010.

[14] F Gugliesi et al., "Nuclear DNA sensor IFI16 as circulating protein in autoimmune diseases is a signal of damage that impairs endothelial cells through high-affinity membrane binding," *PLoS One*, vol. 8, p. e63045, 2013.

[15] S Smith and C Jeffries, "Role of DNA/RNA sensors and contribution to autoimmunity," *Cytokine Growth Factor Rev*, vol. 25, pp. 745-57, 2014.

[16] MH Orzalli, SE Conwell, C Berrios, JA Decaprio, and D Knipe, "Nuclear interferon-inducible protein 16 promotes silencing of herpesviral and transfected DNA," *Proc Natl Acad Sci U S A*, vol. 110, pp. E4492-501, 2013.

[17] L Unterholzner and AG Bowie, "Innate DNA sensing moves to the nucleus," *Cell Host Microbe*, vol. 9, pp. 351-3, 2011.

[18] JD Mcghee, JM Nickol, G Felsenfeld, and DC Rau, "Higher order structure of chromatin: orientation of nucleosomes within the 30 nm chromatin solenoid is independent of species and spacer length," *Cell*, vol. 33, pp. 831-41, 1983.

[19] T Li, J Chen, and IM Cristea, "Human Cytomegalovirus Tegument Protein pUL83 Inhibits IFI16-Mediated DNA Sensing for Immune Evasion," *Cell Host Microbe*, vol. 14, pp. 591-9, 2013.

[20] KM Monroe et al., "IFI16 DNA sensor is required for death of lymphoid CD4 T cells abortively infected with HIV," *Science*, vol. 343, pp. 428-32, 2014.

[21] KE Johnson et al., "IFI16 restricts HSV-1 replication by accumulating on the hsv-1 genome, repressing HSV-1 gene expression, and directly or indirectly modulating histone modifications," *PLoS Pathog*, vol. 10, p. e1004503, 2014.

[22] K Luger, TJ Rechsteiner, and TJ Richmond, "Expression and purification of recombinant histones and nucleosome reconstitution," *Methods Mol Biol*, vol. 119, pp. 1-16, 1999.

[23] MH Orzalli and DM Knipe, "Cellular sensing of viral DNA and viral evasion mechanisms," *Annu Rev Microbiol*, vol. 68, pp. 477-92, 2014.

Chapter 5: Concluding Remarks

The results of the experiments described in this thesis build a current picture of how ALRs recognize DNA, and how this may propagate further downstream to the next components of the signaling pathway. As described in Chapter 2 and Chapter 3, IFI16 and AIM2 both bind dsDNA in a length-dependent manner that shows two-state behavior in solution assays. This binding of DNA is coupled to assembly of a filament along the length of the DNA: disruption of the filament-forming activity results in disruption of DNA-binding, and vice versa. This DNA-binding/self-oligomerization is in large part dictated by the N-terminal pyrin domain, as mutations in the PYD – which makes no interactions with DNA – in turn can lead to disruptions of DNA binding and loss of two-state, cooperative behavior. From these results, a revision of the old model of ALR activation is proposed. The old model put forth a picture of ALRs existing in a state of autoinhibition before binding to DNA; DNA beyond the footprint of the ALR ought to relieve this autoinhibition. As the main candidate for this autoinhibition model is an acidic patch on the face of AIM2-PYD, this was targeted for mutagenesis and found to result not in hyperactivation as would be predicted from an autoinhibition model but in a weakly binding protein that resembled the isolated HIN200 domain under imaging. Length-dependent assembly of a filament upon DNA is a simpler, sufficient explanation of activation.

The length-dependency of DNA binding in turn suggests an answer to the problem of how these sensors distinguish self- from non-self DNA, which was addressed in Chapter 3. As self-DNA is, under normal conditions, sequestered in the nucleus and blocked by various DNA-binding proteins as well as nucleosomes, long

stretches of naked DNA should be relatively uncommon. Invading pathogenic will be exposed upon introduction to the cytoplasm or nucleus before it undergoes transcription or translation, giving a window of time in which these sensors may bind and form filaments upon the naked DNA. A prediction of this model is that DNA protected by DNA-binding proteins – for instance, nucleosomes – should inhibit or prevent ALR oligomerization and DNA binding. This was demonstrated by competition experiments that showed weakened binding to dinucleosomes with varying linker lengths, as well as reduced or abolished time-dependent oligomerization on them. Thus, nucleosomes, abundant on self-DNA, ought to serve as barriers to self-recognition by ALRs.

As to the question of how these sensors feed into the larger pathway of inflammasome activation and innate immune signaling, the results only address one of the two ALRs investigated. For the case of AIM2, the symmetry of the filament matches that of the downstream effector ASC, suggesting that AIM2 may act as a nucleator/template for the polymerization of ASC. This is especially attractive, as similar innate immune systems have demonstrated such behavior, and ASC has been shown to enhance its own oligomerization activity upon introduction of AIM2-DNA complexes. However, there are many outstanding questions in regards to the joining of the sensing portion of the pathway to the later portions. For instance, whether and how IFI16 is able to activate its own downstream effectors remains to be shown, and to date the symmetry of the IFI16 filament has not been determined. Also, whether symmetry is a necessary and/or sufficient means of templating ASC still has to be tested. Thus the results of this thesis raise important questions that

must be tested to determine whether symmetry and templating play large roles, and if so, whether it is a mechanism unique to the AIM2-ASC system or if it is a feature shared by all ALRs.